Hak cipta dan penggunaan kembali:

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

Copyright and reuse:

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

Perpustakaan Univeristas Multimedia Nusantara

BAB II
KERANGKA PEMIKIRAN

2.1 Penelitian Terdahulu

Penelitian ini mengacu pada konsep yang digunakan serta data dan hasil dari beberapa penelitian terdahulu yang terkait jurnalisme data dan *big data*. Salah satu penelitian yang ditemukan adalah jurnal yang ditulis oleh Jan Lauren Boyles dan Eric Meyer yang berjudul "Newsroom Accomodate Data-Based News Work".

Dalam penelitian ini Boyles dan Meyer (2017) melakukan penelitian terhadap 18 pimpinan redaksi koran di Amerika Serikat. Penelitian ini dilatarbelakangi oleh berubahnya sistem pemberitaan di era 2010-an yang mulai bergeser ke berita digital. Beberapa organisasi berita mulai mengintegrasikan jurnalisme data ke dalam produk dan platform jurnalisme. Akan tetapi, untuk mengolah data menjadi konten yang siap untuk dikonsumsi oleh konsumen berita membutuhkan spesialisasi tingkat tinggi. Dibutuhkan keahlian dalam bidang pengkodean dan pemrograman karena untuk menghasilkan jurnalisme data berbeda dengan sistem produksi berita jurnalisme konvensional (Boyles & Meyer, 2017, p. 431).

Tujuan dari penelitian ini adalah mengetahui kondisi redaksi yang menanamkan jurnalisme data dalam organisasinya dan mengetahui struktur organisasi apa yang digunakan untuk mengintegrasikan jurnalis data ke ruang

redaksi (Boyles & Meyer, 2017, p. 432). Hasil dari penelitian ini menghasilkan bahwa ternyata 18 koran di Amerika Serikat ini menghasilkan empat titik kritis, yang pertama, mulanya manajer *newsroom* menganggap bahwa jurnalisme data sebagai sebuah usaha di mana para jurnalis mengambil cerita berbasis data sebagai tambahan saja dari apa yang ditugaskan. Kedua, manajer *newsroom* memperkerjakan seorang praktisi tunggal, yang ditugaskan untuk mengolah data secara *full-time*. Ketiga, memposisikan satu atau dua pekerja data ke dalam struktur tim *newsroom*. Terakhir adalah organisasi berita membentuk divisi data independen yang mana para pekerja beraktivitas secara kolektif sebagai sebuah tim (Boyles & Meyer, 2017, p. 433).

Terdapat persamaan penelitian terdahulu ini dengan penelitian yang ditulis oleh peneliti yaitu sama-sama ingin mencari tahu apa yang membuat suatu *newsroom* menerapkan jurnalisme data dalam pemberitaannya dan mencari tahu mengapa newsroom menggunakan pihak ketiga untuk mengolah data. Selain persamaan, terdapat juga perbedaan dengan yang akan penulis teliti yaitu penulis terfokus akan satu media cetak yang memiliki organisasi pengolah data sendiri namun menggunakan hasil analisis data dari pihak ketiga yaitu *Majalah Tempo*. Selain itu metode yang digunakan pun sama, yaitu studi kasus milik Robert E. Stake.

Untuk penelitian terdahulu yang kedua, penulis menggunakan jurnal yang ditulis oleh Khaerudin Imawan yang berjudul “Jurnalisme Data Resistensi Ruang Siber di Era Post Truth”. Penelitian ini berlatar belakang

dari masuknya ke era digital, kekuatan data sangat dibutuhkan untuk industri manapun, termasuk industri media (Imawan, 2018, p. 1). Dibutuhkan ketelitian dalam mengolah data sebagai sumber berita maupun penulisan berita. Selain itu (Imawan, 2018, p. 6), jurnalisme data juga merupakan salah satu bentuk pemanfaatan *big data* yang dapat dilakukan oleh industri media dan menjadi kebutuhan yang tak bisa dipisahkan dari proses penulisan berita oleh para jurnalis.

Tujuan dari penelitian ini yaitu untuk membedah realitas masyarakat ruang maya dan industri media yang didalamnya terdapat peran jurnalis yang berada di balik data redaksi (Imawan, 2018, p. 1). Persamaan penelitian ini dengan yang akan penulis teliti adalah keduanya sama-sama ingin mencari tahu manfaat data yang diolah membuat apakah disinformasi dapat dikurangi apalagi yang tersebar melalui media sosial. Selain itu penulis juga mengikuti konsep yang digunakan dalam penelitian ini yaitu penggunaan konsep jurnalisme data dan *big data*.

Berikut ini merupakan tabel untuk penelitian terdahulu yang telah dibuat oleh peneliti sebelumnya:

Tabel 2.1 Penelitian Terdahulu

	Peneliti	**Masalah Penelitian**	**Teori dan Konsep**	**Metode**	**Temuan**
1.	*Newsroom Accommo-date Data-Based News*	1. Situasi apa yang membuat organisasi	*Grounded Theory, Investing in Data*	Studi Kasus milik Robert E.	Berkembangnya teknologi digital melahirkan bentuk baru ilmu

	Work (2017) Oleh Jan Lauren Boyles and Eric Meyer	berita menanamkan ilmu jurnalisme data? 2. Apa struktur organisasi yang digunakan untuk menyatukan jurnalisme data ke dalam ruang berita?	*Journalism, Computa-tional Journalism*	Stake, dengan cara wawancara dan menginden ti-fikasi apa yang telah membuat newsroom memprakti kkan jurnalisme data	jurnalistik yang lebih berfokus pada jurnalisme data. Ketika jurnalisme data mulai masuk ke dalam organisasi berita, *newsroom* membutuhkan tambahan orang-orang yang memiliki spesialisasi dalam data. Dikarenakan data merupakan hal yang sangat kompleks dan membutuhkan tim tersendiri dalam *newsroom*. Di sisi lain, jurnalis harus memliki kemampuan untuk menghasilkan data, jika tidak maka jurnalisme data kemungkinan tidak menjadi subbidang yang spesial.
2	*Jurnalisme Data Resistensi Ruang Siber di Era Post Truth*	Bagaimana membedah realitas masyarakat ruang maya dan institusi media yang di dalamnya terdapat peran	1. Media online dan perilaku konsumen, 2. Resistensi jurnalisme data di era post truth, 3. BigData,	*Literature Review*	Di era *post truth*, kebenaran jurnalistik mulai diambil alih oleh berita bohong, berita yang tidak jelas. Solusi bagi krisis *post truth*, yang

	Khaerudin Imawan (2018)	jurnalis yang berada di balik data redaksi	Evolusi Inforgra-fis dan Videogra-fis di Tren Masa Kini		pertama kali dilakukan adalah secara simultan dan akurat mengoreksi kesalahan informasi, dan menunjukkan secara efektif ketidaklogisan dalam berargumentasi untuk menekan praktik disinformasi. Caranya, meningkatkan akurasi produk jurnalisik dengan mempopulerkan Jurnalisme Data. Praktik jurnalistik berbasis *big data* dari sumber yang memiliki kredibilitas tinggi, menganalisanya dengan valid dan menyuguhkan-nya dalam bentuk infografis dan videografi kepada publik.

2.2 Konsep yang Digunakan

2.2.1 Jurnalisme Data

Dalam buku *The Data Journalism Handbook* menyatakan bahwa antara "Data" dan "Jurnalisme" seperti istilah yang sulit. Beberapa orang memiliki pikiran bahwa "data" sebagai sekumpulan angka yang biasanya dikumpulkan di *spreadsheet*. Apa yang membuat jurnalisme data berbeda dari jurnalisme yang lainnya yang merupakan kemungkinan baru untuk menggabungkan kemampuan *nose of news* dan kemampuan untuk membuat cerita dengan menggunakan data di era informasi digital sekarang (Gray et al., 2012, p. 2).

Menurut Bull (2010, p. 323).dalam bukunya yang berjudul *Multimedia Journalism Practical Guide* menjelaskan jika jurnalisme data melibatkan pengambilan sejumlah besar angka dan membuat visualisasi dari ha dapat berupa peta, grafik, atau pun tabel. Visualisasi dari jurnalisme data tersebut untuk mengungkap cerita yang terbungkus dalam data. Jadi, jurnalisme data adalah tentang menggunakan komputer untuk menemukan cerita dari data. Tak hanya itu saja ini tentang menyajikan visualisasi yang dapat dimengerti dengan mudah. Jones dan David (2017, p. 12-18) menjelaskan untuk membuat data dapat disajikan agar mudah untuk dimengerti terdapat tahapan berikut ini:

1. *Getting Data*

   Membahas mengenai data, sebenarnya data itu ada dimana-mana dan banyak hal yang bersangkutan dengan data, bahkan jurnalis dapat menemukan data di mana saja. Namun, layaknya berbagai macam informasi yang jurnalis kumpulkan terdapat beberapa tempat dimana data yang berguna lebih mudah didapatkan. Dimulai dari data yang

dapat dengan mudah diunduh dari situs yang menyajikan *open-data* dan sumber online lain, kemudian beralih ke cara mendapatkan data publik dari pejabat publik, baik secara formal mauapun informal (2017, p. 12).

Namun, dalam hal ini data yang penulis maksud adalah data yang terstruktur, informasi yang terkategorisasi, tersusun di baris dan kolom untuk memudahkan analisis (2017, p. 17). *Open data* telah menjadi salah satu aktivitas besar bagi pemerintah dan sering kali terikat dengan pemerintahan yang terbuka dan transparan. Terdapat lima prinsip yang dibuat oleh pemimpin G8, yang merupakan grup dari negara terindustrialiasasi yang membuat prinsip yang dinamakan *Open Data Charter*, yaitu (2017, p. 18):

a. Pada dasarnya data harus tebuka. Para pemimpin G8 menyetujui bahwa data yang terbuka merupakan hal yang signifikan bagi negara dan masyarakatnya. Meskipun memang beberapa data tidak dapat diakses karena alasan yang sah atas hak milik atau privasi pribadi.
b. Harus ada kualitas dan kuantitas dari data. Ini termasuk komitmen untuk merilis data dari gudangnya (data pemerintah) secepatnya dan memastikan bahwa data tersebut berkualitas.
c. Data harus dapat digunakan oleh semua orang, artinya data itu dirilis dalam format terbuka, dalam volume sebanyak mungkin.

d. Data harus dirilis dengan tujuan adanya tata kelola yang lebih baik. Para pimpinan G8 menyadari bahwa data dapat menguatkan institusi demokrasi dan rintangan administratif

e. Data disajkan untuk inovasi. Semakin banyak masyarakat dan organisasi yang menggunakan data, membuat semakin besar keuntungan bagi sosial ekonomi yang akan dihasilkan.

2. *Analyzing Data*

Ada beberapa poin ketika anda memiliki data, dan yang telah dibersihkan dari salah penulisan, konten yang hilang, dan ketidakkonsistenan lain dari situs *open data*. Bagian ini merupakan penentunya, yang mana menganalisis data untuk menemukan pola dan tren yang akan memberikan pandangan dan cerita (2017, p. 48). Sebelumnya perlu diketahui bahwa di Excel, OpenOffice, LibreOffice, dan program sejenisnya yang mirip dapat dikerjakan dengan *spread sheet* di komputer masing-masing (2017, p. 56).

3. *Visualizing Data*

Ada beberapa software untuk mengolah data dalam bentuk visual, dari Google Maps dan Google Fusion Tables sampai Tableau Public, Many Eyes dan perangkat lunak lainnya. Visualisasi membuat kemungkinan bahwa tidak hanya data saja yang dihasilkan, tapi juga mendeskripsikan cerita dengan data, membuat karakter utama lebih dari sekedar pelaporan berita dengan cara tradisional (2017, p. 165).

Idealnya, visualisasi data harus meningkatkan seseorang mengenai cerita tertentu. Jika kita menerima konsep dasar dari visualisasi data, manfaatnya sangat banyak yaitu tentang seberapa jelas, menyeluruh, dan akurat pencerahannya (2017, p. 167). Visualisasi yang tepat menguatkan, menambah konteks, atau dimensi ekstra ke cerita yang ingin disajikan. Visualisasi yang tidak tepat dapat mengacaukan pesan, atau lebih buruk, mendistorsi atau menyampaikan informasi yang salah. Maka dari itu, dibutuhkan bagan untuk mempermudah khalayak membaca data. Ada tiga jenis bagan yang umumnya dipakai oleh jurnalis yaitu *pie chart*, *bar chart*, dan *line chart*.

Jurnalisme data dapat membantu jurnalis untuk menceritakan hal yang sangat kompleks melalui infografis (Gray et al., 2012, p. 3). Berikut merupakan beberapa alasan mengapa jurnalisme data penting, diantaranya:

1. *Data Journalism Is The Future*

   Jurnalis harus memiliki kecerdasan akan adanya data. Di mana harus meneliti data yang ada dan melengkapi diri dengan berbagai alat untuk menganalisis dan mengambil mana yang menarik. Mengambil salah satu perspektif agar dapat membantu orang memahami dengan apa yang terjadi (Gray et al., 2012, p. 6).

2. *Adapting to Changes in Our Information Environment*

   Berkembangnya teknologi digital membuka jalan baru untuk memproduksi dan menyebarluaskan pengetahuan di masyarakat.

Jurnalisme dapat dipahami sebagai upaya media untuk menyesuaikan dan menaggap perubahan lingkungan informasi, termasuk *storytelling* multidimensi yang interaktif bagi pembaca (Gray et al., 2012, p. 6).

3. *A Way To Tell Richer Stories*

   Dalam dunia data ini dapat muncul ke permukaan untuk membuat suatu cerita lebih komplit dan menjawab pertanyaan dengan cermat dan teliti (Gray et al., 2012, p. 6).

2.2.2 *Big data* Twitter

Menurut Rob Kitchin (dalam Sloan & Anabel, 2017, p. 27) menyatakan bahwa dalam beberapa waktu data itu menampung sejumlah domain yang ada seperti prakiraan cuaca, pasar keuangan, dan lainnya. Seiring dengan perkembangan zaman yang diikuti dengan perkembangan teknologi, dan mulai bermunculan Web 2.0, media sosial, Cloud sebagai penyimpan berkas, dan analitik data yang dirancang untuk mengatasi kelimpahan data. Sejak 2008 dan seterusnya istilah data menjadi daya tarik, apalagi dengan siklus data yang kuat dan banyak yang berpendapat bahwa *big data* ditetapkan untuk merevolusi cara berjalannya bisnis. Tak lama setelah itu, istilah itu beralih ke penggunaan akademis umum disertai dengan argumen bahwa data akan mengubah bagaimana penelitian akan dilakukan ungkap Rob Kitchin (dalam Sloan & Anabel, 2017, p. 28).

Secara sederhana, *big data* adalah sekumpulan data dengan volume yang sangat kompleks untuk dapat diolah dengan teknologi konvensional, selain itu *big data* ini memiliki karakter 3V yaitu *volume, variety,* dan

velocity (Wijaya, 2019, p. 3). *Volume* berkaitan dengan besarnya ukuran data, *variety* menggambarkan banyaknya tipe atau jenis baik dari yang terstruktur hingga yang tidak, dan *velocity* adalah frekuensi munculnya data perubahan data, misalnya tiap detiknya (Wijaya, 2019, p. 17). *Big data* media sosial seringkali digunakan untuk menjawab pertanyaan yang sebelumnya tak pernah sumbernya digunakan. Misalnya, data Twitter yang dapat menjawab mengenai bahasa dari perseorangan atau grup dari wilayah mana dari adanya geotag Twitter (Sloan & Anabel, 2017, p. 34).

Berdasarkan karya Cornelia Brantner dan Jurgen Pfeffer yang berjudul *Content Analysis of Twitter : Big data, Big Studies* dalam *The Routledge Handbook to Developments Digital Journalism Studies* (Eldridge & Franklin, 2019, p. 79) menjelaskan dengan miliaran pengguna dan jutaan cuitan setiap harinya, membuat *big data* dari media sosial telah berhasil menarik perhatian pada ilmuwan sosial. Lebih lanjut penelitian komunikasi di Twitter menjadi istimewa, apalagi di ranah penelitian jurnalisme. Bahwa relevansi Twitter untuk jurnalis dan pemberitaan media mencerminkan betapa pentingnya sumber politik dari informasi dan berita, terutama di *audience* muda.

Merujuk pada sumber yang sama, Twitter dipelajari sebagai salah satu tempat baru sebagai sumber berita bagi jurnalis. Cendekiawan menganalisis bagaimana cuitan ditanamkan dalam reportase berita (2019, p. 80). Studi paling banyak yang disebut adalah analisis media dan jurnalis menangani media sosial dan mengumpulkan cuitan para penggunanya, menggunakan

kualitatif atau kuantitatif ataupun metode digital, termasuk analisis jaringan. Namun ada beberapa yang menggunakan selain pendekatan *automated text-analysis*, seperti *topic modelling*, *sentiment analysis*, atau *machine learning*.

Dalam bukunya, Wijaya memberi contoh implementasi analisa sentimen terhadap cuitan di Twitter. Menurutnya media sosial tidak hanya menjadi urusan bisnis saja, namun sudah secara nyata masuk ke ranah politik. Hasil yang diharapkan dari analisis sentimen ini adalah seberapa besar persentase dari sentimen positif, negatif, dan netral. Dari analisis sentimen, nantinya akan ada beberapa hal yang akan menjadi informasi, yaitu (Wijaya, 2019, p. 150):

a. Persentase sentimen positif, negatif maupun netral dari keseluruhan data yang diolah.
b. Tagar atau *hashtag* apa saja yang berpengaruh signifikan dalam penyebaran sentimen di dalam cuitan para pengguna Twitter
c. *Tweet* atau cuitan apa saja yang mendapat atensi seperti *like*, *re-tweet* dari pengguna Twitter.

Demi mendapatkan hasil analisis data tersebut, dibutuhkan proses untuk memberdayakan *big data* yaitu:

1. *Data Collecting* (Pengumpulan Data)

Pengumpulan Data: *Scraping Twitter* dengan kata kunci agar mendapatkan *tweet* yang mengandung *keyword* tersebut, yaitu dengan struktur data sebagai berikut:

a. *Tweet text* (isi cuitan)

b. *Username* yang mengunggah cuitan tersebut

c. *Mentions* (nama-nama penggguna yang disebiy dalam tiap *tweet*)

d. Tanda pagar (tagar) atau *hashtag* yang dicantumkan dalam tiap cuitan

e. *Retweets* (jumlah *retweet* dari tiap cuitan)

f. Jumlah *like* yang didapatkan dari tiap cuitan

g. Permalink (URI dari tiap *tweet*).

Data cuitan Twitter yang berhasil di-*scrape* dengan kata kunci tertentu biasanya jumlahnya jutaan, maka dari itu disaringlah (*filter*) cuitan tersebut untuk diproses lebih lanjut guna mendapatkan sentimen yang terkandung didalamnya (Wijaya, 2019, p.152).

2. Analisa Data

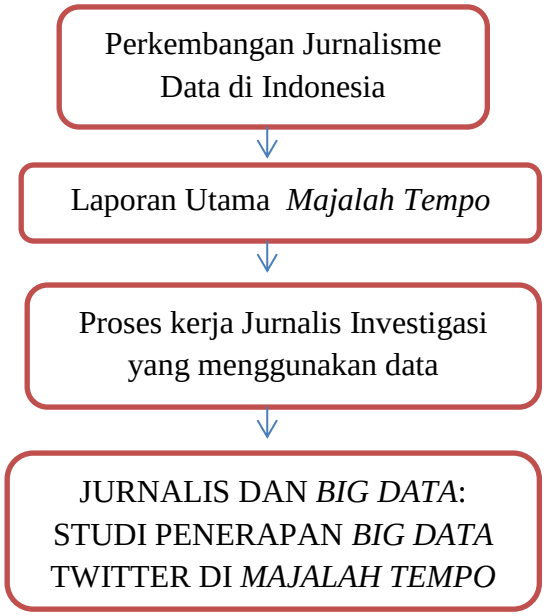
Untuk dapat mengetahui sentimen dari tiap cuitan secara otomatis, tanpa cara konvesional, maksudnya adalah tidak menggunakan tenaga manusia untuk mengolah data, karena membutuhkan waktu yang sangat lama. Maka dari itu dibutuhkan *Machine Learning* untuk keperluan teks. *Machine Learning* adalah cabang dari *Artificial Intelligence* yang fokus pada pengembangan sebuah sistem yang mampu belajar “sendiri” tanpa harus diprogram

oleh manusia berulang kali, yang nantinya melalui sistem ini akan mengeluarkan hasil klasifikasi jenis sentimen (Wijaya, 2019, p.153).

Untuk mengolah *big data* ini diperlukan kemampuan khusus, orang-orang inilah yang disebut *data scientist*. Berdasarkan buku *The Data Science Handbook* dijelaskan bahwa orang yang mampu mengolah data berskala besar biasanya programmer komputer atau ahli dalam machine learning yang ahli dan bekerja di perusahaan *big data* (Cady, 2017, p. 2). Sebenarnya ketrampilan ini, meliputi rekayasa perangkat lunak yang luas dan ekstraksi dari fitur spesifik domain, yang memang pada dasarnya merupakan pekerjaan yang harus diselesaikan oleh *data scientist* (Cady, 2017, p. 3).

2.3 Alur Penelitian

Bagan 2.1 Alur Penelitian



Alur dari penelitian ini pada bagan 2.1 adalah berawal dari laporan utama *Majalah Tempo* edisi 2-8 Januari 2017 dan edisi 4-8 Januari 2018

yang menggunakan hasil analisis *big data* Twitter dalam laporan investigasinya. Peneliti ingin mengetahui bagaimana upaya *Majalah Tempo* dalam memanfaatkan *big data* Twitter dalam peliputannya dan faktor-faktor apa saja yang membuat *Majalah Tempo* bekerja sama dengan pihak ketiga dalam pengolahan data. Dalam penelitian ini, peneliti menggunakan konsep jurnalisme data dan *big data* Twitter.