

Lampiran 1.**Topik Berita Terbanyak Diangkat Selama Januari – Maret 2017**

Topik	Januari	Februari	Maret	Jumlah
Lifestyle	84	92	92	268
Nasional	64	57	79	200
Film	50	34	44	128
Sepak Bola	39	49	37	125
Politik	38	35	36	109
Travel	34	25	50	109
Otomotif	31	36	35	102
Kriminalitas	28	35	29	92
Gadget	25	31	28	84
Karir	28	23	28	79
Kesehatan	41	24	11	76
Internasional	40	14	21	75
Pilkada	5	43	22	70
Kuliner	28	20	15	63
Korupsi	22	13	27	62
Ekonomi	20	19	22	61
Teknologi	23	17	17	57
Saintek	13	18	14	45
Musik	16	12	14	42
Tren	23	7	9	39
Peristiwa	14	4	20	38
Presiden	11	8	12	31
Fashion	0	19	11	30
Hukum	4	10	14	28
Ormas	13	8	7	28
Badminton	1	4	18	23
Infotainment	11	7	3	21
Berita Hoax	10	9	1	20
Balap	4	7	8	19
Bencana	3	8	5	16
Media	2	2	1	5
Seni	1	0	4	5
Quiz	3	0	0	3
Sejarah	3	0	0	3
Militer	1	1	0	2
Olympic	2	0	0	2
Sastra	1	1	0	2

Basket	0	1	0	1
Budaya	0	1	0	1
Opini	1	0	0	1
Redaksi	0	0	1	1
Sosial	1	0	0	1

Lampiran 2.
Jumlah Sebaran Berdasarkan Topik

Topik	Januari	Februari	Maret	Jumlah
Nasional	65019	39438	117595	222052
Kriminalitas	36723	94108	58052	188883
Politik	70188	67971	32383	170542
Sepak Bola	34849	25284	78662	138795
Ormas	67540	30033	36311	133884
Saintek	5303	38101	35235	78639
Peristiwa	20777	3877	43238	67892
Pilkada	7890	42750	8748	59388
Film	22129	2409	26014	50552
Lifestyle	14335	19284	12185	45804
Internasional	24832	11217	4389	40438
Korupsi	13831	6029	19306	39166
Otomotif	3403	7774	22508	33685
Karir	24028	2402	6672	33102
Presiden	14304	2320	13552	30176
Kesehatan	3748	2632	22359	28739
Ekonomi	10262	8029	7723	26014
Teknologi	16567	2376	4100	23043
Travel	5601	8615	5934	20150
Hukum	4406	6976	8392	19774
Bencana	15181	379	609	16169
Badminton	85	819	14256	15160
Kuliner	8228	772	1205	10205
Media	9729	148	73	9950
Sejarah	7460	0	0	7460
Berita Hoax	5388	529	112	6029
Musik	1387	1580	1182	4149
Seni	14	0	4052	4066
Gadget	603	2313	1145	4061
Tren	1369	620	1023	3012
Balap	1020	133	1791	2944
Infotainment	2046	569	80	2695
Quiz	1054	0	0	1054
Fashion	0	655	210	865
Sastra	74	57	0	131
Sosial	127	0	0	127
Militer	79	14	0	93

Budaya	0	59	0	59
Opini	50	0	0	50
Redaksi	0	0	26	26
Olympic	4	0	0	4
Basket	0	1	0	1

Lampiran 3.**Jumlah Topik Berita Jurnalisme Paling Banyak Diangkat Selama Januari – Maret 2017**

Topik	Januari	Februari	Maret	Jumlah
Nasional	14	9	14	37
Ekonomi	7	5	12	24
Korupsi	12	0	6	18
Pilkada	3	10	3	16
Otomotif	3	4	3	10
Kesehatan	6	2	1	9
Politik	4	2	1	7
Kriminalitas	2	1	2	5
Film	4	0	1	5
Lifestyle	0	1	4	5
Internasional	5	0	0	5
Bencana	1	2	1	4
Teknologi	1	1	1	3
Infotainment	1	0	1	2
Travel	1	0	1	2
Ormas	1	0	1	2
Peristiwa	2	0	0	2
Media	1	0	0	1
Presiden	1	0	0	1
Musik	1	0	0	1
Karir	0	1	0	1
Sepak Bola	0	1	0	1
Berita Hoax	0	1	0	1
Gadget	0	1	0	1

Lampiran 4.**Jumlah Sumber Data yang Digunakan Selama Januari – Maret 2017**

Sumber Data	Jumlah
BPS	36
KPU	19
AIMSI	12
Kemenkes	10
APJII	6
Beritagar	5
GIKBI	5
Pemprov DKI	5
Transparency International	5
WHO	5
World Bank	4
cdc.gov	4
Kepolisian Rep Indo	4
Kementerian PAN RB	4
Kementerian Keuangan	4
filmindonesia	3
Cinema 21	3
UC Web	3
Kemendikbud	3
Kementerian Perindustrian	3
filmindonesia	3
Other	3
ICW	3
Kementerian Pariwisata	2
Ombudsman Republic Indonesia	2
Cisco	2
Kementerian ESDM	2
Setara Institute	2
PIHPS	2
hargapangan.id	2
BPBD	2
Bank Indonesia	2
SMRC	2
BNN	2
BPPSPAM	2
SKK Migas	2

KPK	2
ICJR	2
UNDP	2
BKPM	1
YLKI	1
The Future of Chinese Travel	1
World Economic Forum	1
Kemnaker	1
OAG Punctuality Justice	1
United Nations University	1
Bawaslu	1
Global Slavery Index	1
GHI	1
Fronterra Brands Indonesia	1
BMKG	1
Job Planet	1
FAO	1
Konsorsarium Pem Agraria	1
PLN	1
KLHK	1
GISMI	1
Gaikindo	1
Kemendagri	1
UEFA	1
Kemenkominfo	1
BKPM	1
Pricewater Cooper	1
Gartner	1
Indo Barometer	1
Bloomberg	1
Kepolisian Jawa Barat	1
APNBP	1
Wahid Institute	1
Serious Fraud Office	1
Kementerian Hukum & HAM	1
Spotify	1
British Phonotic Industry	1
Dirjen Penyediaan Perumahan	1
BPIH	1
Riset Kesehatan Dasar	1
CNN	1

Independent	1
Data.go.id	1
University Georgia	1
DJPS	1
IDC	1
KNKT	1
Dirjen Pajak	1
IIF	1
PPATK	1
PNBP	1
Google	1
The Wrap	1
Global Corruption Barometer	1
PPIM	1

Lampiran 5. Interview Guide

INTERVIEW GUIDE

Dalam upaya mendapatkan data penelitian tentang jurnalisme data dalam portal berita online Beritagar.id. *Interview guide* berikut ini akan menjadi pedoman dalam wawancara sebagai metode penelitian yang digunakan dalam penelitian ini.

A. Gambaran Umum Jurnalisme Data

Pertanyaan berikut ini bertujuan untuk menggali informasi tentang gambaran umum jurnalisme data secara teoritis:

1. Menurut Beritagar.id, apa itu jurnalisme data? Apa yang membedakan jurnalisme data dengan jurnalisme yang lain?
2. Bagaimana penerapan jurnalisme data di dalam redaksi Beritagar.id?
3. Menurut David Pearce Demers dan Suzanne Nichols (1987), jurnalisme data dapat diibaratkan sebagai sains karena juga menggunakan metode penelitian *social science* dalam praktek jurnalisme. Bagaimana pendapat anda tentang hal tersebut?
4. Apa itu data? Ada berapa jenis data yang ada, dan apa saja?
5. Seberapa besar peran data dalam jurnalisme data? Data seperti apa yang bisa digunakan dalam jurnalisme data?
6. Bagaimana cara menyeimbangkan antara data sebagai sumber informasi dengan sumber-sumber lain (wawancara, observasi, dokumen, dll)?
7. Jurnalisme data identik dengan praktek jurnalisme yang menggunakan teknologi. Teknologi seperti apa yang digunakan oleh Beritagar.id (Machine Learning, Natural Language Processing, Computer Assisted Reporting)?
8. Apakah jurnalisme data juga bisa dipraktekkan tanpa perlu teknologi yang canggih? Seperti apa contoh jurnalisme data tanpa teknologi canggih?
9. Bagaimana peran jurnalisme data dalam praktek jurnalisme masa kini? Apa tujuan di balik penggunaan jurnalisme data?
10. Sumber informasi jurnalisme data lebih bersifat akurat dan komprehensif (lengkap). Bagaimana pengaruh jurnalisme data dalam memberantas berita bohong (*hoax*) yang menjadi salah satu masalah besar di era media baru ini?

11. Apakah ada hubungan antara jurnalisme data dengan kualitas berita? Seberapa besar pengaruh jurnalisme data dalam meningkatkan kualitas sebuah berita?

B. Jurnalisme Data dalam Beritagar.id

1. Dalam Beritagar.id, rubrik apa saja dan konten yang seperti apa yang paling sering menggunakan jurnalisme data?
2. Seperti apa proses penulisan jurnalisme data yang digunakan Beritagar.id? Siapa saja yang bertanggung jawab bagi masing-masing proses tersebut? Seperti apa tugas masing-masing penanggung jawab proses jurnalisme data?
3. Beritagar.id memiliki 2 fitur media berbasis data, Lokadata dan Berkas. Apa perbedaan antara Lokadata dan Berkas (*living data*)? Apa tujuan dari masing-masing fitur tersebut?

C. Indikator Proses Jurnalisme Data

Menurut Philip Meyer (2002), indikator-indikator dalam proses jurnalisme data meliputi 6 proses, diantaranya adalah pengumpulan data (*collecting*), penyimpanan (*storing*), pengambilan kembali (*retrieving*), analisa (*analyzing*), pengurangan (*reducing*), dan penyampaian data (*communicating*). Pertanyaan berikut ini bertujuan untuk menggali informasi tentang indikator-indikator proses jurnalisme data tersebut.

1. Dari mana saja data bisa didapatkan? Kalau data didapatkan dari lembaga tertentu, apakah data yang didapatkan bersifat mentah atau sudah diproses? Bagaimana lisensi atau hak cipta dalam menggunakan data tersebut? Apakah data tersebut berbayar atau tidak? Data yang seperti apa yang berbayar dan tidak berbayar?
2. Untuk data yang dikumpulkan sendiri, bagaimana Beritagar.id mengumpulkan data-data tersebut? Menggunakan metode pengumpulan data yang seperti apa saja? Pengumpulan data yang seperti apa yang paling sering digunakan?
3. Berapa lama waktu yang dibutuhkan untuk mengumpulkan data dalam satu proyek? Siapa saja yang bertugas dalam mengumpulkan data tersebut?
4. Bagaimana cara menentukan dokumen apa saja atau siapa saja yang digunakan sebagai informan dalam pengumpulan data? Seperti apa bentuk data mentah yang didapatkan saat mengumpulkan data pertama kali?

5. Dimana data-data tersebut disimpan? Bagaimana cara menyimpan data-data tersebut? Apakah data disimpan dalam jangka waktu tertentu, selamanya, atau dihapus hanya pada kondisi tertentu? Seberapa besar ukuran/kapasitas data dalam satu proyek data?
6. Apakah ada aturan atau syarat tertentu dalam menyimpan data?
7. Dalam proses *retrieving*, data biasanya didapatkan dari pihak ketiga. Bagaimana Beritagar.id memilah data yang digunakan agar sesuai dengan kebutuhan berita?
8. Bagaimana cara mengolah data agar data tersebut dapat bercerita dengan sendirinya tanpa perlu ditambah narasi? Lalu bagaimana cara menambahkan narasi agar tidak terkesan berlebihan (karena sudah ada data)?
9. Bagaimana cara menentukan penting atau tidaknya data tertentu? Lalu, apa yang dilakukan pada data yang tidak terpakai?
10. Apa saja yang dilakukan saat menganalisis data agar tercipta berita dengan kaidah jurnalisme data yang komprehensif tapi juga menarik bagi audiens?
11. Data dipresentasikan atau disampaikan dengan apa (statistik, grafik, tabel)? Bagaimana cara membuatnya? Menggunakan aplikasi apa?
12. Bagaimana cara mempresentasikan data yang menarik namun tetap sesuai dengan topik yang ditentukan sejak awal?

D. Lain - Lain

1. Apakah jurnalisme data tetap bisa dilakukan oleh orang biasa tanpa menggunakan teknologi atau teknik yang rumit? Aplikasi apa yang bisa digunakan (secara gratis maupun berbayar) untuk dapat membuat jurnalisme data secara individual?
2. Apa saja aspek-aspek penting yang harus diperhatikan dalam pembuatan jurnalisme data?
3. Apakah jurnalisme data memiliki pengaruh yang penting dalam memberantas berita bohong (*hoax*)? Seberapa penting jurnalisme data harus dilakukan?
4. Apakah semua media lintas harus menggunakan jurnalisme data agar berita yang disajikan lebih akurat? Bagaimana cara menggerakkan media-media agar sadar akan pentingnya memberantas berita bohong?
5. Apakah ada hubungannya antara jurnalisme data dengan berita yang berkualitas?

6. Apa kritik dan saran bagi dunia jurnalistik di media Indonesia? Apakah media di Indonesia telah memenuhi kaidah jurnalistik, khususnya yang berkaitan dengan berita bohong atau berita yang tidak akurat?
7. Apakah kritik dan saran bagi peneliti atau akademika lain yang juga hendak melakukan penelitian tentang jurnalisme data?

Lampiran 6. Transkrip Wawancara

Transkrip Wawancara

Narasumber 1:

Aghnia Adzkia (Aghnia) – *Data Journalist* Beritagar.id

Peneliti : Nama saya Galang dari UMY. Akan melakukan wawancara untuk penelitian skripsi yang berjudul *Jurnalisme Data* dalam beritagar.id. Mungkin pertanyaan-pertanyaan selanjutnya akan berkaitan dengan jurnalisme data. Entah secara konseptual atau praktis yang diterapkan di beritagar. Sebelumnya terima kasih juga sudah diberi kesempatan buat melakukan wawancara. Langsung saja ke pertanyaan pertama. Pertanyaan awal masih secara konseptual. Menurut mbak, jurnalisme data itu seperti apa? Dan apa yang membedakan antara jurnalisme data dan jurnalisme lainnya?

Aghnia : Ini aku bicara sebagai wartawan berita di beritagar dan bukan mengatasnamakan beritagar sendiri. Jadi, bisa saja apa yang aku sampaikan itu berbeda dengan kebijakan kantor atau bagaimana cara kantor memandang jurnalisme data. Itu yang harus di *clarified*. Terus, menurut aku jurnalisme data itu salah satu tren jurnalistik. Bagaimana cara kita *story telling* atau cara kita menceritakan sebuah berita itu berdasarkan data. Data itu kan biasanya kebanyakan orang melihat data kualitatif. Entah itu kutipan dari narasumber, hasil *press release*, atau itu dari cerita di *report* atau di laporan gitu. Tapi kalau yang sekarang ini agak bergeser sedikit ke data statistik. Jadi bagaimana sebuah berita itu. Jadi jurnalisme data itu bercerita menggunakan data statistik dan cerita itu muncul dari si data statistik. Sebenarnya ada berbagai macam cara orang memandang atau *technically implementation* dari jurnalisme data itu. Data itu sebagai komplemen dari data kualitatif yang ada. Misalnya diselipkan jadi dua pragraf satu berita. Sebenarnya cerita itu bukan dari data itu sendiri. Terus yang kedua, menurut aku data itu sebagai *lead*-nya atau judul dari artikel itu sendiri. Kebayang gak sebelumnya? Kayak udah pernah membaca artikel data itu enggak?

- Peneliti : Yang diselipkan itu ada kan mbak? Ada berita-berita yang cuma data aja. Jadi kayak cuma disajikan data terus kita yang menginterpretasikan dan menarasikan sendiri. Itu seperti apa yang kayak gitu mbak?
- Aghnia : Nah itu menurut aku salah satu produk jurnalisme data. Itu di Indonesia jarang yang menerapkan itu paling katadata dan beritagar itu juga beberapa kali seperti itu. Kalau di beritagar sendiri itu ada yang namanya rubrik artikel data. Itu dari (rupa data). Itu kayaknya dulu 5 kali seminggu. Terus sekarang udah mulai, mulai minggu ini atau minggu lalu, itu 3 kali seminggu. Itu membahas data. Jadi kalau misalkan, udah pernah baca belum?
- Peneliti : Sudah.
- Aghnia : Kayak gitu kan maksudnya yang kamu tanyain tadi? Itu menarasikan data. Terus itu satu hal implementasi jurnalisme data dari beritagar. Contoh lainnya lagi, setiap artikel harian kayak misalkan artikel harian apapun. Misalkan soal *Asian Games* pengen tahu bagaimana mereka *support* penonton yang difabel. Apakah tempat lokasinya untuk menonton atau tempat lokasi turnamennya itu ramah difabel? Bisa buat pengguna kursi roda? Bisa untuk yang buta dan segala macamnya? Data di situ, di artikel seperti itu *supporting* nya itu dalam bentuk misal jumlah penderita difabel di Jakarta. Jadi data itu sendiri bukan jadi *main story*, *lead*, *highlight*, atau judul, tapi dia hanya jadi pelengkap dari satu cerita utuh yang narasinya itu diambil dari data kualitatif.
- Peneliti : Jadi bisa dijadikan sumber informasi utama atau sebagai pelengkap aja ya?
- Aghnia : Iya, betul seperti itu.
- Peneliti : Lalu untuk data itu sendiri sebenarnya ada jenis-jenisnya enggak sih? Lalu fungsinya data itu apa? Kayak misalnya jenis macam-macam data. Kalau misalnya cara penyampaiannya kan ada yang secara statistika, grafis, tabel. Nah dari data itu ada macam-macamnya enggak? Atau mungkin didefinisikan aja data itu apa?
- Aghnia : Kalau data yang konteksnya jurnalisme data, itu spesifik ke data kuantitatif, data angka, data (...), kalau tren yang berkembang sekarang ya di seluruh dunia. Terus kalau jenis data kalau aku sendiri melihat ada dua jenis data. Data deskriptif dan data

eksploratif. Data deskriptif itu data yang kalau misalkan cuma menunjukkan narasi deskripsi aja. Misalkan jumlah penderita difabel di Jakarta. “Oh ada sekian yang cacat fisik, oh ada sekian yang cacat mental, oh ada sekian yang menggunakan kursi roda”. Tapi kalau data eksploratif itu akan lebih dari itu. Kita enggak hanya melihat jumlah difabel di Jakarta, tapi kita melihat misalkan *layer* yang lain lagi. Mereka itu ada di wilayah mana aja. Terus mereka itu berangkat dari latar belakang keluarga yang seperti apa? Bagaimana pendidikan mereka? Itu data eksploratif. Artinya dia enggak cuma satu *incomer* aja, tapi dia bisa beragam indikator dan dari situ dicari eksplor lainnya. Misal oh ketika dilihat sama lokasi. Oh paling banyak di Jakarta itu penderita cacat mental atau kejiwaannya terganggu. Itu misalkan paling banyak di Jakarta Barat dan mereka itu, aku enggak tau ini asumsi ya. Misalnya mereka itu pendidikan paling tingginya SMA. Itu kan udah lebih ke eksplor, bukan cuma sekedar ini jumlahnya sekian. Nah yang menarik itu bagaimana kita bisa menemukan satu cerita dari data itu. Jadi kita bisa menarasikan dan mengeksplorasi jadi satu artikel atau satu berita. Misal data eksploratif itu aku dulu pernah mengolah (...*dating*) data terorisme di Indonesia. Ada beberapa kejadian teror di Indonesia dari tahun 1977-2018. Nah, ketika kita cuma bilang jumlahnya aja, itu akan jadi deskripsi dasar. Itu nanti juga bersinggungan dengan bagaimana cara menyampaikan ya. Tapi ketika kita belah misalkan di Indonesia sendiri. Sebelum 2007 dan sesudah 2007, itu akan kelihatan beda. Misalkan sebelum 2007, terus dilayer juga sama sasaran teroris. Sebelum 2007 itu sasaran terorisme itu masyarakat sipil yang *western* atau yang non-Islam. Dan setelah 2007, sasaran mereka itu beralih ke polisi. Nah itu kan enggak cuma sekedar jumlah atau deskripsi yang aku bilang dari awal. Tapi itu dia mengeksplor dia bisa cerita, dia bisa dapat *angle* yang baru bahwa “Oh sebelum 2007 ternyata sasarannya paling banyak masyarakat sipil. Setelah 2007 paling banyak adalah polisi. “Oh kenapa sih kayak gitu?” Ternyata terjadi perubahan tren siapa kelompok radikal yang berkuasa di Indonesia pada periode tersebut. Misal pada tahun sebelum 2007 itu yang berkuasa adalah Jamaah Islamiyah. Setelah 2007 yang berkuasa adalah Jamaah Ansharut Daulah. Mereka itu punya paham yang berbeda. Jadi ketika mereka menyerang dengan bom atau tembakan, itu sasarannya berbeda. Karena Jamaah Ansharut Daulah, yang setelah 2007 itu, mereka lebih menganggap polisi itu sebagai musuh

utama, daripada masyarakat sipil atau non-Islam yang bule-bule itu juga. Kamu masih inget enggak dulu tahun 2001, ada kayak bom bali? Itu masih Jamaah Islamiyah. Nah, Jamaah Islamiyah itu menganggap masyarakat sipil bule-bule di Bali itu sebagai musuh dari Islam. Nah, kalau yang setelah 2007, musuh Islam itu bukan mereka tapi polisi. Makannya serangan-serangan kayak Thamrin, bom yang di Sarinah itu yang diserang kan polisi. Terus yang di Surabaya yang terakhir, yang diserang juga ada kantor polisi. Jamaah Ansharut Daulah mengaku bahwa polisi itu adalah musuh dari Islam karena polisi menolong pemerintah yang kafir di Indonesia. Dari satu data itu dia bisa cerita banyak banget hal. Nah itu data eksplor. Data yang bersifat eksploratif. Jadi dia bisa punya satu cerita sendiri dan dia dari situ bisa dicari “Oh penyebabnya apa? Oh kenapa bisa jadi seperti ini?” Yang deskripsi tadi cuma bilang “Oh, jumlah terorisme meningkat di Indonesia. Eh maaf, jumlah kejadian teror di Indonesia itu fluktuatif dan yang paling tinggi itu tahun 2001 dan 2002. “Oh pada saat itu ada bom Bali”. Se-*scientific* itu kalau data deskripsi. Kalau data eksploratif dia bercerita, di *-layer* lagi dengan cerita *why*. Biasanya kalau data deskripsi itu menjawab kalau enggak 5W1H. Kalau data deskripsi itu biasanya mereka menjawab *what*, *when*, terus *who*, tapi kalau di data eksplor itu biasanya menjawab *why* dan *how*. *How* jarang banget sih, tapi seenggaknya dia muncul, yang paling sering *why* dan ada tren apa ini.

Peneliti : Jadi misalnya data di dalam artikel dituliskan berdasarkan data BPS bahwa jumlah terorisme segini-segini cuma dalam bentuk tulisan aja, itu termasuk jurnalisme data juga?

Aghnia : Itu termasuk jurnalisme data yang alirannya deskriptif dan dia komplimen. Jadi dia pelengkap doang. Dia bukan jadi suatu narasi utama.

Peneliti : Jurnalisme data kan identik dengan praktik jurnalisme yang menggunakan teknologi. Nah, khusus di beritagar ini menggunakan teknologi apa aja? Lalu penggunaannya seperti apa?

Aghnia : Jelas kita pakai teknologi banget. Beritagar itu juga ada pakai robot. Pernah tahu enggak Robotornalnya beritagar?

Peneliti : Belum.

- Aghnia : Robotorial ini, jadi kita udah enggak pakai manusia sama sekali. Jadi algoritmanya bekerja sendiri dan dia *publish* berita.
- Peneliti : Mulai dari pengumpulan datanya?
- Aghnia : Iya, mulai dari pengumpulan datanya. Itu pakai mesin. Yang ngebangun adalah (..... produk). Itu dia cerita bisa bola dan saham.
- Peneliti : Yang pengumpulan datanya bisa langsung diambil secara *online* ya mbak?
- Aghnia : Iya dan dia langsung nih kayak gini. Ini salah satu bentuk jurnalisme data. Misalnya ini robot yang nulis.
- Peneliti : Ini yang nulis artikelnya robot?
- Aghnia : Iya, ini yang nulis artikel robot dan ini salah satu bentuk jurnalisme data, karena dia menjelaskan data bagaimana. Nah ini robot yang ngerjain. Kalau kamu nanya soal penggunaan teknologi di dalam jurnalisme data, ada levelnya macam-macam. Kalau di beritagar yang kamu maksud jurnalisme data yang dikerjakan robot, itu 100% penggunaannya pakai teknologi, pakai mesin, pakai algoritma di balik itu. Nah, kalau yang aku kerjakan adalah yang kayak di Loka Data itu. Yang mungkin kamu belum punya gambaran soal robotorial. Nah itu nanti jadi temuan menarik, kamu eksplor lagi di sana. Ada level penggunaan teknologi. Nah ini yang ngurus aku, yang Loka Data itu. Nah ini dari data awalnya. Kalau ini penggunaan teknologi tuh kayak mungkin bisa dibilang 50:50 sama manusia. Artinya aku enggak bisa mengerjakan ini kalau tanpa bantuan teknologi komputer. Mungkin aku bisa mengerjakan ini tapi akan sangat lama prosesnya. Dengan teknologi tuh mempercepat banget buat menyelesaikan ini dalam waktu satu hari atau dua hari. Sampai mulai dari *scrapping*, ngolah data sampai visualisasi. Mungkin kalau tanpa teknologi bisa 3 hari sampai seminggu. Teknologi ini levelnya sama seperti *human*. Jadi dia membantu tapi dia enggak sepenuhnya teknologi kayak robotorial. Karena di sini aku harus mengerjakan laporan. Jadi prosesnya ini, aku dapat data, aku olah, aku wawancara orang hasil dari temuanku. Ternyata sinkron dan jadilah satu berita. Kalau di robotorial, dia enggak perlu wawancara dan dia 100% komputer dan teknologi. Jadi dia cuma ngumpulin data yang terbuka. Udah ada data saham, data sepak bola. Entar kamu lihat robotorial yang

sepak bola. Nah itu mereka ngumpulin semua trus habis itu mereka *publish* sendiri. Mereka bikin teks sendiri. *Promote*-nya tentu di *treat* sama kita ya. Maksudnya itu kayak IT *developer*, bikin algoritmanya dan bikin korpus-korpus katanya supaya jadi kayak gini tuh mereka. Kalau aku tadi harus wawancara, harus ngolah, aku harus cari *angle* yang menarik apa nih? Kamu *familiar* kan sama *angle*? Kira-kira apa nih yang menarik? Dari sudut pandang mana nih yang menarik. Robot bisa punya kapasitas yang sama seperti manusia ketika mereka melihat *angle*. Mangkannya kenapa, data yang bisa diolah sama robot, itu data-data yang belum pasti. Itu data-data yang berpola. Misalnya IHSG, terus data sepak bola. Itu bisa diolah sama robot. Data gempa bumi, data cuaca, itu data-data yang sifatnya sudah *ready* dan polanya terbaca dan itu bisa dilah sama si robot. Tapi kalau misalnya kayak (...) gitu yang masih butuh bantuan manusia, itu kebutuhan untuk teknologinya 50%.

Peneliti : Kalau berita-berita, data-data yang kita harus mengambil ceritanya sendiri itu enggak bisa dikerjakan di robotorial ya?

Aghnia : Iya, karena robotorial itu yaudah gitu. Kalau kamu baca berita-berita robotorial, kamu akan melihat “Oh monoton banget gayanya”. Gayanya akan sama, karena *angle*-nya akan sama. Lihat aja dari judulnya yang tiga artikel terakhir. Sama kan? Ini kelihatan banget kalau yang ngerjain robot. Meskipun tulisannya mungkin enggak kelihatan kayak robot. Tapi kalau kamu baca ya gitu. Itu misalnya yang bola.

Peneliti : Kalau untuk program yang digunakan untuk robotorial itu tahu enggak mbaknya?

Aghnia : Itu namanya *machine learning*. Bagian dari *machine learning* itu *Natural Language Processing*. Itu tadi sama kang deny. Kalau aku akan menjawab yang tiga kali seminggu itu.

Peneliti : Nah untuk yang 50:50 itu yang *human* juga, itu aplikasi yang dipakai apa?

Aghnia : Kalau aku cerita dari awal ya. Kalau misalkan proses mengolah data mulai dari pengambilan data sama penulisan sampai visualisasi itu ada banyak tahapan. Biasanya yang aku pakai buat ambil data itu *Google Spreadsheet*, sama *most of the time google*

spreadsheet, kayak 90% dari pekerjaan aku, aku pake itu, *scraping* sama olah data. Pernah enggak ya aku pakai *software*? *Software* pun aku pake (exchone) yang ada di *google spread sheet*. Mengolah pake *google spread sheet most of the time*. Mengolah kadang juga pake *Tableau* sama visual. *Tableau* itu kayak misalnya kita *upload* data itu bisa *drag down*. Itu buat lihat paling tinggi yang mana dan yang paling tinggi yang mana. Kalau visualisasi kadang pake *Tableau* atau pake ini, jadi kita punya satu data *hub* gitu. Namanya *lokadata.beritagar.id*, mesin ini dibuat sendiri. Kalau ini enggak pakai (per part), kalau *Tableau* kan pakai (perpart) kayak gini-gini. Ini bagian dari *beritagar* juga. Yang kayak gini ini mesinnya main sendiri. Kata Data juga punya kayak gini. Jadi bangun sendiri sistemnya di *back end* gitu. Jadi kita tinggal *upload* data trus ini bisa dimasukkin.

Peneliti : Berarti ini yang bikin grafik ini loka data *programmer* sendiri atau pakai aplikasi lain?

Aghnia : *Programmer* di dalam *beritagar* membuat *back end* seperti ini. Orang-orang data, dan wartawan yang bisa data kayak aku gini, bisa mengakses ini dan bisa *upload* data di sini untuk nanti dikirimkan ke teman-teman redaksi ketika menulis. Biasanya artikel harian yang butuh cepet terus hais itu kayak butuhnya misal dalam waktu satu jam dua jam itu pakai ini biasanya. Karena ini sistemnya sudah ada tinggal kita *upload* terus tinggal di *embed*. Nah gitu. Tapi kalau misalkan yang mingguan biasanya pakai visual statis. Kayak ini, *image* nah pengolahan gambarnya ini pakai *Tableau* itu terus tinggal dipoles sama tim grafis. Terus kalau yang harian itu adalah ini yang harga telur dan ayam. Ini pakai yang sistem dibangun sendiri tadi. Ini harian, butuh cepet. Jadi enggak perlu grafik yang bagus dan warna-warni, yang penting dia kelihatan langsung.

Peneliti : Nah untuk menemukan berita-berita yang data itu, sebenarnya saya agak kesusahan. Soalnya di Loka Datanya tidak mencantumkan berita yang ke situnya kan mbak. Jadi kadang saya harus nyari satu-satu mana aja yang pakai loka data. Saya buka satu-satu. Kalau buka langsung di loka datanya cuma datanya aja mbak ya?

Aghnia : Iya. Kalau misalnya gitu, kamu lihat di *beritagar* lalu dia ada setelah editorial itu loka data. Ini dapatnya tiga kali seminggu,

bukan yang setiap hari ada. Ini yang mingguan. Ini lebih panjang. Nah kalau yang jenis loka data, yang tiga kali seminggu ini, dia eksploratif. Kalau yang harian, itu yang deskriptif yang cuma jadi pelengkap. Tetapi kadang-kadang kebanyakan yang artikel loka data biasanya cuma deksriptif sih. Bergantung bagaimana cara penyampaiannya. Deskriptif dan eksploratif enggak cuma data itu sendiri, tapi bagaimana cara menyampaikan. Ada lagi enggak?

Peneliti : Terus rubrik apa aja sih yang biasanya pakai jurnalisme data atau berita-berita dengan tema apa yang paling sering menggunakan data?

Aghnia : Kalau berita dengan tema apa yang paling sering menggunakan data, aku enggak punya datanya, tapi secara kasar berita ekonomi itu selalu ada datanya dan bisa *disupport* dengan data. Karena *somehow* mereka ngomongin angka. Terus ngomongin data itu pasti sering banget *disupport* harga telur naik, inflasi, realisasi pendapatan DKI Jakarta. Terus misalkan ada saham *freeport* pasti *disupport*. Terus berita-berita yang jarang pakai data itu berita-berita, yang kayak apa ya. Karena rata-rata semua pakai data.

Peneliti : Iya rata-rata pakai data atau studi atau *research*.

Aghnia : Tapi sebenarnya data itu enggak melulu kayak yang harus serius, ekonomi gitu enggak. Kita pernah ngeluarin data yang *fun*. Kayak misalnya Slank, soto, terus sambel, terus ada Dilan. Pernah baca Dilan enggak? Nah, ini Dilan aku yang ngolah datanya sama dibantu satu orang, sama aku yang nulis.

Peneliti : Kalau yang Slank itu kayaknya eksploratif banget ya mbak?

Aghnia : Sebenarnya kalau Slank itu deskriptif. Dia ada angka dan dia akan jelasin gitu doang. Itu deskriptif sih. Nah yang ini berita yang aku tulis, tentang rayuannya Dilan. Jadi terus dibandingkan sama, *Ada Apa dengan Cinta? Dan Galih dan Ratna*. Ini pakai *Tableau*. Coba kamu klik ini, terus nanti akan kelihatan. Jadi ketika si Dilan itu ke-PD-an lagi merayu Milea, trus Milea itu responnya apa. Itu di sini. Jadi kelihatan. Terus begitu juga sebaliknya. Gitu. Dan ada film, lagu, terus aku pernah buat Indonesia Idol. Pernah nulis berita Indonesian Idol. Lagu 20 kontestan Indonesian Idol. Jadi misalkan kalau kontestannya perempuan, dia sering nyanyiin lagu cewek atau cowok sih. Terus ini juga, perjalanan 20 kontestan.

- Peneliti : Berarti yang untuk ini pengumpulan data primer ya mbak? Maksudnya yang benar-benar dikumpulkan sendiri. Jadi mbak Nia mendengarkan 20 lagu itu sendiri?
- Aghnia : Iya. Enggak cuma 20, ya semua. Jadi kalau kamu tanya data primer dan sekunder itu masuk ke cara kita ngambil data. Ada data yang *available*, ada data yang mesti *coding* sendiri, mesti ambil sendiri dan yang primer itu kayak Indonesian Idol, kayak Dilan, kayak Slank. Ada data-data yang sudah kita pakainya data *release* dari pemerintah atau dari lembaga (...) atau dari lembaga riset gitu. Nah itu *i can't see* data sekunder ya. Terus kita mau *review* kalau bahasanya deskripsi.
- Peneliti : Kalau data sekunder itu berbayar apa enggak?
- Aghnia : Nah itu pertanyaan baik. Karena kalau dari BPS itu kita punya *micro data*. Nah itu kita beli produknya itu di level terkecil dari sebuah kuesioner. Misal ada pertanyaan kuesioner X gitu, kita bisa tahu si orang itu jawab Ya atau Tidak. Misalnya jawab poin-poin itu. Kalau itu *micro data* itu beli kalau dari BPS. Tapi kalau kita mau data yang secara *general* aja enggak usah beli pun ada dan banyak banget dari BPS *nge-release* itu berdasarkan jenis surveynya. Misalnya survey Sosial Ekonomi Nasional, itu setahun ada dua kali, september dan maret. Mereka pun juga bikin publikasinya dan itu lengkap banget. Macem-macem, coba kamu cek sendiri. Ada impor dan ekspor, kayak gitu juga.
- Peneliti : Berarti sebagian besar lembaga pemerintah itu sudah mendukung *open government data* atau belum mbak?
- Aghnia : Aku belum bisa bilang itu, udah bisa mendukung *open government data* karena beberapa kasus aku temuin di *website*. Hal yang sederhana, misalnya kita butuh data itu kan mesti melihat *website* mereka. Ada beberapa lembaga kementerian yang *website*-nya enggak *accessible*, terutama buat orang-orang kayak aku wartawan yang untuk berita. Misal Kementerian Koperasi dan UMKM. Kemarin aku dapat data dari BPS jumlah koperasi dan jumlah volume usaha atau *cash flow* yang ada di koperasi tersebut di seluruh Indonesia. Aku cuma pengen dapat data yang tahun 2017 aja, itu susah banget. Aku cuma dapet publikasi dari BPS cuma sampai tahun 2016, karena memang belum jadwal rilis untuk tahun berikutnya. Terus aku cobalah ke *website* Kementerian

Koperasi dan UMKM itu, terus mereka enggak ada datanya. Terus aku coba telepon ke bagian data mereka. Telepon kantor. Juga ribet banget, lama banget, dan aku minta tolong ke *email* aku sore ini, dan mereka menyanggupi itu. Tapi sampai 3 hari enggak dikirim-kirim data. Padahal aku cuma butuh data jumlah koperasi dan volume usaha tahun 2017, agregat seluruh Indonesia. Terus begitu, ada temen redaksi lainnya mau nulis tentang koperasi, terus mereka mau wawancara petinggi koperasi, barulah petinggi itu yang ngasih intruksi buat ngasih data. Jadi kayak sistem birokrasi yang sangat ribet dan tidak terbuka, itu membuat aku dan semua data analiser itu kesulitan untuk menemukan data. *I can't see* kalau misalkan itu sudah mendukung atau belum. *It's not point yes*. Banyak sudah mulai terbuka, enggak seperti itu. Tapi ada beberapa hal kayak gitu itu masih ditemui. Misalnya aku mengolah data tentang adat seluruh Indonesia, perhutanan adat. Mungkin kamu bisa *Google on letter*. Aku data kayak gitu enggak ada di Kementrian Lingkungan Hidup. Datanya semua lama dan aku harus ketemu orangnya dulu. Minta datanya dulu dan baru dikasih. Kalau enggak kayak gitu mereka enggak akan *update* di *website*. Tapi ada beberapa kementrian yang *update*. Contohnya BPS sih yang paling bagus. Karena mereka badan survey. Kalau mereka enggak bagus malah jadi pertanyaan.

Peneliti : Koinfo dan Kementrian Kesehatan gimana?

Aghnia : Kalau Kementrian Kesehatan biasanya mereka rilis publikasi gitu. Kayak laporan Surat Kesehatan Dasar setiap 5 tahunan. Mereka memang buka sih datanya. Jakarta termasuk wilayah yang *open government*-nya lebih baik daripada wilayah lain. Tapi bukan itu bagus dari yang paling bagus tapi lebih baik. Dalam artian ketika aku cari data penyandang difabel tahun 2015. Aku dapat datanya. Tetapi datanya masih tahun 2015. Padahal aku pengen data yang sudah sampai 2017. Terus kadang dapat data yang itu enggak sinkron penamaannya. Itu enggak bisa kita *track* data *update*-nya atau memang belum dimunculin di *website* mereka. Mendukung atau enggak, lebih baik dari sebelumnya, tapi enggak sepenuhnya sudah mendukung. Masih perlu banyak PR untuk beberapa kementrian tentang *open data*-nya. Mungkin mereka ada ya, cuma mereka itu enggak *accessible*.

Peneliti : Terus seberapa penting sih peran *open government data*?

Aghnia : Penting bangetlah. Kalau kita bicara soal Indonesia, Kementerian Pendidikan Indonesia itu bagus. Ketika kita bicara soal angka partisipasi sekolah atau angka buta huruf di Indonesia. Enggak mungkin dong nyari ambil data di *World Bank*. Mungkin bisa, karena *database*-nya bagus banget. Tapi *World Bank* dapetnya dari Indonesia kan? Kenapa enggak aku cari yang primernya dulu? *World Bank* itu dapetnya dari BPS atau Kementerian Pendidikan. BPS itu dapetnya Kementerian Pendidikan. Daripada aku ambil di sana, mendingan aku ke Kemendikbud. Kemendikbud ini salah satu yang bagus. Jadi seberapa penting *open goverment*? Penting banget. Karena kalau aku mau bicara data Indonesia, ya aku harus dapat dari Indonesia dulu. Enggak mungkin aku dapat dari *World Bank* atau *International Institution*. Toh mereka juga minta ke pemerintah Indonesia. Kalau kasus Kementerian Pendidikan, itu salah satu kementerian yang bagus. Mereka selalu bikin publikasi dan bikin *database* juga itu lengkap banget dan selalu *update* dan aku gampang banget dapat data angka buta huruf, angka partisipasi sekolah, sarana yang rusak di seluruh Indonesia atau provinsi.

Peneliti : Untuk membuat data primer yang Slank atau Dilan, Indonesian idol itu biasanya durasinya berapa lama pembuatannya?

Aghnia : Kalau data primer ya? Kalau Slank aku enggak tahu, karena itu dibuat sebelum aku gabung di beritagar. Slank itu juga proyek bulanan atau tiga bulanan. Jadi dia *micro side*. Cuma kalau yang Indonesian Idol atau yang soto, itu butuh waktu yang lebih lama. Daripada kita sekedar mengolah data yang biasa. Mungkin teman-teman ngebayanginnya “Kalau tema-tema yang seru kayak Indonesian Idol, Soto, sambel, itu datanya gampanglah, dan termasuk cepet ngerjainnya”. *No*. Malah justru tema-tema yang *fun* dan ringan gitu susah banget ngolah datanya dan itu butuh waktu sampai 3 hari.

Peneliti : Karena harus benar-benar satu-satu ya mbak?

Aghnia : Iya, karena kita harus ngumpulkan satu-satu dan kita kayak memilah satu-satu. Sementara kalau data-data yang keras, dalam artian data-data yang sudah ada. Misalnya kredit bunga tiap bank selama 1 tahun ini triwulannan. Itu sudah dapet gampang banget dari BI untuk tahun 2016 doang atau dari OJK tahun 2017. Tapi kalau Indonesian Idol, seberapa banyak kontestan indonesia idol yang nyanyi lagu-lagu asing. Seberapa banyak konstestan cewek

yang nyanyi lagu-lagu cowok. Itu kan kayak enggak mungkin ada datanya itu (*pages*) *like* itu tuh enggak gampang ngumpulannya dan itu tuh butuh kesabaran dan ketelitian.

Peneliti : Khusus di beritagar sendiri, untuk di redaksi beritagar itu yang membuat jurnalisme data itu bagian apa aja? Lalu tiap bagian-bagian itu tugasnya apa aja?

Aghnia : Jadi, redaksi itu ada berapa ya. Cuma kalau tim data itu sekitar ada, biasanya kalau artikel mingguan itu dikerjakan oleh 3 orang. Jadi yang mengolah, menulis dan visual. Itu minimal 3 orang, yang ngolah ini kebanyakan adalah tim data. Kadang aku bantu meskipun aku anak redaksi. Aku bantu temen-temen data untuk ngolah. Terus yang nulis itu tim redaksi, terus yang visual itu tim visual. Gimana bisa caranya yang ngolah data itu orang lain, tapi yang menuliskan orang lain. Padahal orang lain ini juga harus tahu cara baca datanya segala macam. Biasanya ada di rapat redaksi. Rapat redaksi itu penentuan temanya. Di situ penentuan tema, digodoklah datanya apa aja yang ada. Kalau udah menyamakan persepsi, tinggal nanti tim data mengolah dan selama proses pengolahan itu mereka juga komunikasi sama tim redaksi. Ini lho aku menemukan data yang ini. Terus tim data itu memberikan catatan kecil, nanti catatn itu yang diolah oleh penulis. Misalkan dari catatan ini bisa *follow up* ke narasumber yang kompeten untuk menjawab pertanyaan. Contohnya, soal berita fasilitas tenda di Mina jadi sorotan jamaah haji yang baru rilis kemarin. Ini awalnya adalah temen data mengolah indeks kepuasan haji tahun 2017. Mereka mengolah, terus aku tanya temuan apa yang menarik. Ternyata tenda di Mina, salah satu wilayah yang bersebelahan dengan Mekkah di Arab gitu. Itu pelayanan tenda di sana itu buruk banget. Pelayanannya itu dapat skor paling rendah dibandingkan indikator yang lain di indeks tersebut. Paling rendah itu 70,26 atau berapa gitu. Nah anak data bilang gitu. Ini yang menarik. Pelayanan tenda ini, nilainya paling rendah. Terus indeksnya menurun dari tahun sebelumnya. Terus aku wawancaralah ke orang sambil aku riset. Penulis, aku biasanya riset. Kenapa sih tenda di Mina itu pelayanannya amenurun. Kata Jamaah haji, ternyata di Mina itu luasnya sedikit, tapi jumlah jamaah haji itu meningkat terus setiap tahun. Misal dari tahun 2016 dan 2017 ada peningkatan 30%. Sementara segitu doang, jamah hajinya banyak, otomatis kan jadi berdesakan, dari situlah aku

dapat stori. Jadi ketika temen data kasih tahu unsur *what* dari sebuah tulisan, aku menjawab *why* dan *how*. Semua itu proses dilalui dengan komunikasi dua arah. Sejak dari pembahasan tema. Kayak gitu.

Peneliti :Berarti prosesnya penentuan tema, cari data atau gimana?

Aghnia : Biasanya aku dan teman-teman data mencari teman lewat data. Jadi riset kecil-kecilan. Kira-kira data apaan nih yang menarik buat dijadikan tulisan. *Preliminary research* ini diperlukan untuk menentukan atau mengajukan tema ketika rapat redaksi ketika hari jumat setiap minggunya jam 4 sore. *Preliminary research* untuk data ini, untuk nyari *angle*. Tapi emang enggak bisa dipungkiri, kadang ketika berdiskusi atau rapat sama redaksi dan data terus kok enggak cocok. Kayak misal *story*-nya enggak dapat. Aku sebelumnya sudah meriset kurs rupiah terhadap dolar Amerika dari tahun 1950-an sampai 2017. Aku usulkanlah ini ketika rapat. Ini ada data menarik di mana peningkatan kurs rupiah dari yang zaman dulu tahun 1950-an itu cuma berapa rupiah, enggak lebih dari Rp 10 terus jadi Rp 250 puluhan tahun selanjutnya. Terus sekarang jadi Rp 14.000. Terus tim redaksi bilang “Wah enggak bisa nih kalau cuma ini doang. Kan ada inflasi, jadi enggak *make sense* ketika kita cuma ngangkat kurs doang”. Terus aku berpikirlah lagi. Terus ada salah satu usulan di redaksi. “Orang enggak akan peduli, yang mereka pedulikan adalah dengan uang segini mereka bisa beli berapa barang?”. Terus aku berpikir “Gimana kalau misalkan kita bandingkan kurs Indonesia terhadap dolar Amerika di tahun, misal setelah reformasi 1998-2018 dan pada saat itu 1 dollar bisa beli berapa gram emas. Dengan diskusi seperti itu, akhirnya terbentuklah satu *angle* untuk satu data. Jadi *preliminary research* yang aku lakukan adalah aku mengumpulkan data. Kalau aku enggak sanggup mengumpulkan data karena waktu yang singkat aku akan cuma baca-baca data aja. Jadi sebelum maju ke teman-teman redaksi dan data untuk rapat hari jumat itu, aku kayak udah riset-riset. Nah setelah rapat dimatengin temanya. Habis itu setelah rapat itu disetujui, akhirnya teman-teman data ngolah data. Memang beberapa kali, kadang yang akan kita *preliminary research* terus sudah dimatengin di sini, kadang semuanya beda ketika kita akan mengolah datanya. Dan kalau udah kayak gitu harus ada *adjustment* misalnya ganti *angle*, ganti bahasa atau datanya ditambahin.

- Peneliti : Kalau bicara tentang *storing* sama *big data* nih untuk beritagar sendiri, gimana cara menyimpan datanya. Maksudnya lebih ke data-data mentahnya.
- Aghnia : Di *cloud*. Tapi kalau loka data itu dia pakai sistem namanya C-Kam. Jadi yang loka data beritagar itu dia pakai sistem c-Kam. Yang kamu tanyain soal tutorial, tentang *machine learning*, (*NLP*) sama C-Kam. Sama untuk yang artikel 3 kali seminggu sama buat yang artikel *support* harian di *cloud*, *Google Spreadsheet* di *Google Drive*. Untuk *security*-nya alhamdulillah *secure*. Selama ini enggak ada masalah. Toh itu juga yang kita ambil adalah *open data*. Itu sudah ada dan itu tinggal mengolah. Kalau data primer, toh itu data primer bukan yang data *clasify*.
- Peneliti : Kalau mengambil jurnalisme data dengan mengolah dari data mentah, itu gimana sih caranya untuk menemukan cerita atau menemukan *angle* dari data mentah atau mungkin data-data yang diambil dari data sekunder?
- Aghnia : Itu butuh *skills*. Karena enggak semua orang lihat data langung “Oh ini menarik”. Kadang *reference* atau konteks itu tidak terlihat. Kalau analisis wacana ada teks dan konteks. Kalau data itu kadang cuma data aja, teks aja. Nah konteks yang menjelaskan data ini itu yang butuh tahu. Kayak misalkan aku lagi mengolah *series* Indonesia Merdeka. Lagi-lagi kurs rupiah terhadap dolar dan data itu sendiri harus disertai dengan konteks. Misal aku lagi mengolah *series* Indonesia Merdeka. Ada bagaimana kurs dolar terhadap Amerika. Ada terjadi kenaikan. Selama 1953-2018 itu udah terjadi apa aja sih di Indonesia. Untuk menemukan *angle* cerita, kita harus riset. Itu kalau kita udah riset, kita akan dapat konteksnya dan *story*-nya. Misalnya tahun 1979 terjadi krisis energi. Pada tahun 1990-an terjadi inflasi sampai 600%. Oh misalkan di tahun 2008 terjadi krisis global. Konteks-konteks cerita yang didapat dari riset itu akan memperkaya si data dan kita bisa dapat *angle* atau *story*.
- Peneliti : Nah agar di proses akhir pembuatan jurnalisme data di visualisasi gimana caranya untuk mempresentasikan data, entah itu menggunakan statistika atau grafik biar lebih menarik atau menggunakan aplikasi apa?
- Aghnia : Kalau aplikasi kayak yang tadi aku cerita kebanyakan pakai *Tableau*. Kalau misalkan untuk artikel yang 3 kali seminggu itu

setelah pakai *Tableu* distatiskan jadi dibikin kayak *image* gitu. Terus cara mempresentasikan data itu, kadang-kadang datanya banyak banget. Tapi bagian mana sih yang relevan sama cerita? Kayak ini nih, yang tenda di sini ya. Nah kan aku mau cerita sebenarnya banyak banget nih. Indeks kekuasaan haji. Di dalam indeks kepuasan haji itu ada kategorinya. Ada buslah, bus antarkota, ada tenda, ada segala macam. Nah, karena yang menarik itu adalah pelayanan tenda di Mina, makannya itu yang jadi *highlight* dari visualisasi. Kalau itu jadi *highlight* dari visualisasi karena itu adalah kategori terendah, makannya yang paling cocok untuk ditelusuri adalah pakai diagram atau grafik batang kayak gini. Jadi yang pertama kali dilihat dari visualisasi pasti kayak apa sih yang mau kamu ceritain, tulis, apa yang mau kamu tampilkan di dalam visual. Jadi orang ketika melihat visual tuh jadi ini maksudnya. Tapi itu butuh *skills* juga. Kadang aku ketika dapat satu *story* aku bingung (fis)nya. Yang Dilan itu aku harus diskusi sama Pak Dodi dia *supervisor* aku di data. *Supervisor* aku ada dua. Aku harus diskusi sama dia banget. Kayak gimana sih caranya membuat orang itu, ini *user friendly*. Lalu data apa yang mau aku munculin, itu *skills* juga sih. Kayak misalnya aku mau nulis kurs, itu enggak mungkin aku tunjukkan dalam grafik pie. Itu enggak *make sense*, makannya aku tunjukkin dalam bentuk *line* atau grafik batang. Yang paling *matching*. Karena beberapa grafik pun juga. Ada beberapa data yang itu cuma bisa ditampilkan dalam bentuk beberapa grafik. Enggak semua grafik bisa ditampilkan. Kayak *time series*, misalkan. Data historis dari tahun 1960 atau dari mana sampai sekarang. Itu lebih cocok untuk ditampilkan dalam grafik *line* atau grafik batang. Kalau grafik pie yang lingkaran itu, karena grafik pie itu lebih cocok misal persentase. Misalkan persentase penduduk yang menggunakan bahan bakar untuk memasak. Misal di tahun 1950-an itu cuma 78 keluarga Indonesia yang menggunakan kayu. 80% menggunakan kayu, 20% menggunakan minyak tanah. Itu akan dilihat lebih mudah itu kalau misalkan pakai grafik batang.

Peneliti : Pertanyaan terakhir, untuk memilah data gimana sih cara memilih antara data yang penting sama data yang sebenarnya enggak perlu dimasukkan itu enggak apa-apa?

Aghnia : Itu masuk di dalam proses yang namanya *data cleaning*. Kan udah ngumpul, kita bersihkan mana yang kita butuhkan dan mana

yang tidak kita butuhkan. Misal pengeluaran setiap individu untuk misalkan beli makanan sama *traveling* atau misal sama beli baju segala macamnya. Kita udah *scraping* dapat data banyak banget. Mulai dari persentasenya, berapa jumlahnya, nominalnya, berapa rupiah untuk beli makanan. Makanannya makanan jadi atau minuman jadi. Misal beli (padi) atau beli apa. Kalau kita mau ngambil cerita “Oh lebih banyak pengeluaran untuk makanan dibandingkan baju misalkan. Itu kalau misalkan dikeluarkan angka aslinya itu bisa. Jadi persentasenya dihapus atau diilangin. Tapi kalau kita mau membahas pola pembelian makanan atau baju, berarti angka ini udah enggak penting. Kita cuma butuh persentase untuk setiap sub itu dalam beberapa tahun. Jadi gimana caranya untuk memilah data yang kita butuhkan atau enggak itu bergantung kita mau cari apanya.

Peneliti : Berarti balik lagi ke *angle*?

Aghnia : Iya balik lagi. Misal dulu saya mau mengolah data diabetes dan jumlah kalori yang dikonsumsi oleh setiap orang Indonesia pada tahun 2017. Setelah *discrapping* datanya gak *make sense*. Diabetes enggak serta merta digabungkan dengan kalori. Karena yang menjadikan diabetes itu kan kalori yang tersisa. Sementara kalau asupan itu kalori yang dimakan. Kita enggak ada data kalori yang terbakar. Akhirnya itu dibuang. Kita akhirnya coba cari *angle* baru dan itu temanya bener-bener baru. Akhirnya ambil obesitas. Itu pun juga enggak langsung digabungkan dengan kalori, tapi *di-layer* dengan yang lain. Jadi untuk memilah data mana yang ternyata enggak penting, bergantung kita mau nulis apa dan seberapa relevan data itu untuk ditulis. Ketika itu sudah tidak relevan sama sekali ya dibuang aja cari yang baru.

Narasumber 2:

Rahadian Prajna Paramita (Rahadian) – Wakil Pemimpin Redaksi Beritagar.id

Peneliti : Mas Dian, sebelumnya perkenalkan saya Galang dari mahasiswa Ilmu Komunikasi Jogja, UMY. Sedang melakukan skripsi dengan mengambil tema tentang jurnalisme data di beritagar. Saya ada beberapa pertanyaan-pertanyaan tentang jurnalisme data secara konseptual dan bagaimana jurnalisme data diterapkan di beritagar. Jadi saya akan tanya ke mas Dion tentang lebih ke secara teknisnya gimana jurnalisme itu diterapkan di beritagar. Pertanyaan mungkin masih secara konseptual dulu. Menurut mas Dion, atau menurut beritagar itu jurnalisme data itu apa? Apa yang membedakan jurnalisme data dengan jurnalisme yang lain?

Rahadian : Kalau kita sih aslinya kan *data driven*. Karena kita *data driven* maka *story*-nya itu memang di-*drive* oleh data. Itu yang akan membedakan ketika ada *story* yang ditempel data. Kalau ada peristiwa pembunuhan diliput jadi berita, beritanya adalah berita pembunuhan. Kalau di dalamnya nanti diberi data seberapa banyak orang yang sudah terbunuh, itu lain. Kalau kita melihat *data driven*, yang lihat berita, *story* itu bukan peristiwa tetapi data. Bukan peristiwa seperti jurnalisme pada umumnya. Reportase itu kan peristiwa. Kalau *data driven* basisnya yang men-*trigger story* itu data bukan peristiwa. Kalau ada peristiwa sebagai (*respect*) itu lain. Kita dua-duanya kita pakai. Ada *story* dengan peristiwa yang dilengkapi data, ada yang benar-benar kita ngejelasinnya pakai data, di-*drive* oleh data.

Peneliti : Kalau secara definitif data itu apa sih mas? Lalu fungsi data sebagai apa?

Rahadian : Dalam data jurnalisme yang disebut data itu selalu kuantitatif. Mau kualitatif yang dikuantitatifkan enggak masalah. Tapi dasarnya selalu angka-angka statistik. Jadi mengkuantifikasi semua hal. Jadi kalau peristiwa itu adalah informasi kualitatif, kalau kita agregat ya jadi kuantitatif. Ya itu data. Data itu bisa diolah jadi informasi. Jadi data peristiwa yang selalu basisnya data-data yang sifatnya, yang bentuknya yang kita maksud dalam (.....). Selalu begitu. Tapi kan bentuk data macam-macam kualitatif juga

bisa. Tapi data peristiwa yang dimaksud dalam *data journalism* itu statistik.

Peneliti : Lalu seberapa penting perannya jurnalisme data dalam praktek jurnalisme? Lalu kenapa beritanya lebih menonjolkan data (...) jurnalismenya?

Rahadian : Karakteristiknya yang paling menarik dari data *driven*-nya itu kan *scientific*. Ada yang menyebutkan sebagai *precision journalism*. Kata itu yang sebenarnya dipakai sebagai rujukan juga sebagai mereka menyebutnya sebagai jurnalisme yang presisi. Itu juga relatif, makudnya dalam tingkat tertentu yang namanya akurasi itu tetap mengandung *human error*. Semua survey mengandung *human error*. Presisi itu relatif sih. Tidak bicara pada level akurasi 90%-100%, pasti ada *miss*-nya. Kita tahu *link*-nya sampai seperti apa. Menariknya adalah dia cenderung datar. Maksudnya gini, kalau melaporkan hari ini ada pembunuhan di angkot, itu kan terjadi di satu waktu tertentu, di satu wilayah tertentu. Kadang-kadang dalam pemberitaan itu membuat generalisasi. “Oh terjadi penusukan penumpang angkot di Bekasi”. Terus tiba-tiba Bekasi dilabel sebagai kota yang tidak aman. Bukan salah beritanya, tapi wujud berita itu membuat orang mengambil kesimpulan yang tidak profesional. Dengan menggunakan *data driven* kita bisa ngasih tahu. Ya mungkin peristiwa itu betul sadis, tetapi tidak terjadi setiap hari juga. Jadi kita bisa ngomong Bekasi masih aman. Karena statistik mengatakan, tidak serta merta jadi generalisasi yang tidak proporsional tadi. Kalau mau ngomong jurnalisisasi yang proporsional ya mainnya dengan data. Kita mencari *angle* lain yang berbicara tentang bagaimana sih status kriminalitas kita dalam 5 tahun terakhir . Baru kita bisa mengambil kesimpulan oh bekasi paling tidak aman. Paling tidak aman itu pembandingnya apa? 5 tahun yang lalu atau dengan wilayah lain, itu kan beda lagi. Kalau kita ngomong soal historikal nya bekasi ya kita lihat 20 tahun atau 5 tahun terakhir ada kemacetan enggak? Spesifik tentang penusukan di angkot misalnya. Mungkin terjadi setahun berapa kali gitu. Tapi karena penusukannya membuat orang mati itu bisa menjadi *story*. Misalnya pelakunya ibu hamil. Itu akan menjadi level kesadisannya ini kan kualitatifkan. *History*-kan cenderung begitu. Data bisa masuk untuk melengkapi, untuk memberi perspektif yang lebih profesional. Ya jangan nge-judge ini sadis, ya enggak proporsional juga. Bahwa satu orang mati itu

problem. Tapi tidak berarti orang harus bereaksi secara berlebihan. Kecuali (...) untuk mendukung situasi yang buruk itu, ya tugas media untuk nge-post pemerintah untuk melakukan sesuatu. Ini bukan kasus yang pertama kali, atau kasus yang jarang terjadi. Ini udah rutin, sebulan sekali ada. Itukan sudah *something*. Kok sudah bertahun-tahun, 5 tahun terakhir, setiap bulan terjadi, ini ada apa? Harusnya kan pemerintahkan ambil langkah.

Peneliti : Lalu jurnalisme data kan identik dengan *scientific* atau teknologi. Nah khususnya di beritagar ini, beritagar itu menggunakan teknologi apa aja sih mas? Kalau saya dengar ada *machinel learning* dan *natural language processing*. Bisa dijelaskan dikit enggak mas?

Rahadian : Kamu pernah baca sejarahnya *CAR* enggak? Computer Assisted Reporting itu sebenarnya udah ada sejak dulu. Itu salah satu yang dipakai untuk bicara mengenai sejarahnya *New York Time* mengolah data lewat *spreadsheet*. Disitu muncul kesadaran, ternyata komputer ini bisa membantu kita untuk (...). karena pengolahan *spreadsheet* kita jadi tahu berapa persen bisa (mematikan...) itu teknologi. Dengan data yang sekarang sudah semakin banyak, teknologinya jadi semakin relevan. Karena misalnya tadi saya baru buka di datanya Amerika itu ada catatan pelanggaran lalu lintas setiap hari selama berapa tahun. Pokoknya total datanya itu tabel teks (...) yang besarnya hampir 600 mega. Itu kalau enggak ada teknologi setengah mati mau nyiapin gimana, besar sekali. Kalau kita ngomong tabel itu *row*-nya udah berapa berapa juta, kolomnya udah berapa puluh. Jadi matriks yang luar biasa besar yang dikerjakan secara manual itu tidak bisa lagi. Jadi butuh teknologi. Teknologi yang relevan apa aja yang bisa mengolah data itu. Tinggal bergantung jenis datanya apa. Kayak angka-angka itu tadi kan itu (...). Karena banyak sekali sekarang teknologi yang bisa dipakai mengolah. Kalau cuma untuk sekian juta baris sekarang bisa dilakukan dalam hitungan detik. Yang rutin kita lakukan itu mengolah data BPS. Kita kan belinya data micro, bukan data publikasi yang sudah muncul di depan. Data micro itu isinya jawaban kuesioner. Kalau ada 600.000 responden ya 600.000 kali baris. Itu jutaan kolom. Sedangkan kita mau narik misalnya berapa belanja keluarga untuk keperluan rumah tangga. Itu ada di (...) kolom. Itu kan sebenarnya mau pilih kolom ini sama dari *row* yang mana. Misalnya kita pilah perprovinsi dan per

kabupaten kota. Semua bisa dilakukan dalam hitungan detik. Ada banyak teknologi sekarang yang *support* untuk kayak gitu. Kalau *machine learning* itu sudah tahap yang layak, karena penerapan *machine learning* ini kan sebenarnya dipakai untuk hal-hal yang bisa (di...) oleh mesin. Yang kita praktekin misalnya kalau sekarang itu kalau mau bikin artikel-artikel yang sekarang basisnya robot. Itu sekarang semua sudah diproses secara otomatis. Kita udah bikin *template*. Kalau angkanya naik, kalimat ini. Kalau angkanya turun pakai kalimat ini. Jadi itu sudah disusun sendiri oleh parameter yang sudah kami tetapkan. Kita mau pakai parameter yang mana. Nah itu dia mencari sendiri. Istilahnya sudah *rulesbased*. Selama *rules*-nya terpenuhi, syaratnya terpenuhi akan memunculkan ini dan seterusnya. Itu yang paling gampang. Sebenarnya itu juga basisnya *mostly* data. Kalau kita (...) saham itu sebenarnya isinya angka juga. Akhirnya kalau turun, yang naik siapa aja, berapa persen. Itu sebenarnya bagaimana data diotomasi menjadi narasi.

Peneliti : Kalau untuk yang robotorial itu berarti semua berita dikerjakan oleh robot ya? Nah program yang digunakan apa aja sih mas? Lalu fungsi masing-masing program itu untuk apa?

Rahadian : Itu basisnya cuma bahasa pemrograman. Bahasa pemrogramannya hampir sama aja. Karena bukan aplikasi ya, itu bahasa pemrograman. Itu general. Misalnya mau pakai (Python mau pakai PHP) terserah. Sebenarnya sama tapi yang kita pakai basisnya PHP kalau enggak yang Python itu yang simpel.

Peneliti : Lalu secara redaksi untuk secara redaksi untuk penulisan jurnalisme data, itu tim yang berperan siapa aja?

Rahadian : Sebenarnya kita cuma punya dua tim. Redaksi yang didukung oleh tim data. Semua artikel itu ya kolaborasi antara dua tim itu. Jadi ada yang ngolah data, diskusi sama yang nulis. Jadi tidak kayak bayi berjalan gitu. Bisa selesai di sini, di sini. Prosesnya kan atraktif bisa bolak-balik. Dari pengolahan data dievaluasi, “Kayaknya ini ada *angle* lain deh”. Terus direvisi lagi sama penulisnya, penulisnya bisa lihat. Bisa berubah di tengah jalan. Karena data ini ketika diolah kadang tidak sesuai dengan bayangan awalnya.

- Peneliti : Terus biasanya data itu dikumpulkan dari mana aja sih mas? Kayak data primer dan sekunder itu. Atau mungkin data-data yang dikumpulkan sendiri, itu mengumpulannya bagaimana?
- Rahadian : Kalau kita memang mengandalkan data-data panjang, nyaris tidak mengerjakan sendiri. Maksudnya data BPS itu salah satu sumber terbesar. Karena kaya datanya. Yang lain itu belum ada yang bisa sebanyak dan seluas BPS. Pemerintah lain mengeluarkan data kadang-kadang lengkap, kadang-kadang enggak. Jadi tidak rutin. Beberapa ada yang rutin. Kayak Kemenkumham itu punya laporan jumlah napi, setiap saat *update*. Itu bisa diambil datanya, bisa dipakai. Kalau mau cerita soal kepadatan lapas. Itu datanya di situ semua. Padatnya berapa persen, *over capacity*-nya berapa persen. Itu bisa dipakai dan *update*-nya bagus. Yang kita lakukan sendiri itu enggak terlalu banyak. Biasanya ya kayak film, Dilan itu bisa dicari sendiri. Filmnya ditonton, dicatet, bisa manual. Kalau kita dengerin Slank ya lagunya didengerin, manual. Karena itu butuh waktu, butuh orang kita tidak terlalu sering begitu. Enggak terlalu rutin, hanya untuk selingan. Paling rutin ya lembaga yang *credible*. Kayak misalnya (...) yang merilis data yang memang diakui kredibilitasnya.
- Peneliti : Lalu proses jurnalisme datanya yang mulai dari pengumpulan, *cleaning*, atau *scrapping* itu di sini prosesnya seperti apa mas?
- Rahadian : Ya pola itu kan *planning*, itu sudah standarnya begitu. Semua data mentah itu harus *dicleaning* dulu. Kalau sudah *dicleaning*, kalau sudah diolah nanti ketahuan *angle*-nya apa nanti divisualisasikan. Kalau sudah divisualisasikan ya tinggal dimasukkan ke artikelnya menjadi bagian dari tulisan. Cuma gitu aja.
- Peneliti : Biasanya kalau untuk *cleaning* atau *scrapping* itu pakai aplikasi apa sih mas untuk memvisualisasikan? Aplikasi yang digunakan apa aja?
- Rahadian : Paling *basic* ya *spreadsheet*. Mau pakai *excel* boleh, mau pakai *google sheet* boleh, mau pakai *open office* juga boleh. Tapi semua *basic spreadsheet* itu bisa digunakan sebagai proses *cleaning*. Karena matrikulasi data itu selalu dalam bentuk tabel dan kolom. Jadi semua programnya ya rata-rata di *spreadsheet* yang dasar. *Cleaning*-nya di situ. Visualisasi bisa. Kita punya (*dashboard*)

visualisasi sendiri. Sekarang (...) peta. Sebagian besar sekarang artikel itu juga ngambil dari situ. Tinggal masukkan tabel, keluar visualisasi grafiknya. Kemudian ada yang dibikin statik, *image*. Nah itu ada *designer*-nya yang mengolah data. Tapi biasanya dirancang lewat *Tableau*. Kalau kita melihat datanya harus interaktif dan rumit, ya *Tableau*-nya yang kita pakai, artikel. Tidak dibuat statik dulu. Karena keuntungan data ada yang harus interaktif, karena datanya terlalu rumit. Nah itu dicover sama statik ya udah. Biasanya *Tableau* yang dipakai.

Peneliti : Lalu data-data yang sudah digunakan itu disimpan di mana? Lalu *big data* di beritagar seperti apa?

Rahadian : Untuk *storage* juga pakai sistem. Karena seluruh data yang kita pakai itu sudah ada klasifikasinya. Jadi ada pusat datanya, *database* untuk data itu ada yang dicatat ada yang mengendalikan. Sudah tersistem sih. Jadi bisa *diupdate*. Kemarin kasus terorisme, kalau ada data tahun lalu ya datanya *diupdate*. Kayak data-data kejuaraan bulu tangkis, itu (...). Ya tinggal hasil turnamen terakhir tanggal berapa, *diupdate* terus. *Google* kayak gini semua kan mudah diakses. Tinggal aksesnya aja, siapa yang lihat, siapa yang bisa edit. Pokoknya semua sudah dikatalogkan. Kita sebutnya katalog data.

Peneliti : Untuk pertanyaan selanjutnya mungkin agak di luar teknikal sama redaksional. Mungkin untuk orang-orang biasa yang tidak punya akses untuk menggunakan teknologi yang canggih dalam pembuatan jurnalisme data itu apakah mereka juga bisa mengaplikasikan, menerapkan jurnalisme data juga meskipun tanpa teknologi?

Rahadian : Harusnya bisa. Teknologi yang susah, menurut saya, maksudnya itu teknologi yang dasar. Yang paling penting memang bukan teknologinya, yang paling penting orangnya. Karena ya itu hanya *tools*. *Tools* bisa *powerful* kalau anda tahu cara pakainya gimana. Kalau *excel* itu punya kemampuan yang sebenarnya lumayan *powerful*. Anda mau statistik seperti apa bisa diolah di situ. Pokoknya unsur-unsur dasar statistiklah. Mau nyari korelasi dan seterusnya itu sebenarnya ada di situ semua. Jadi relatif. Cuma karena mungkin orangnya aja yang enggak punya kebutuhan kesitu atau karena memang enggak pernah pakai. Selama dia bisa paham

statistik dasar untuk digunakan dalam penulisan ya harusnya enggak terlalu susah.

Peneliti : Berarti pada umumnya pembuatan jurnalisme data itu aksesnya masih terbuka ya mas ya?

Rahadian : Harusnya iya. Yang jadi masalah kan pengetahuan.

Peneliti : Lalu apa sih mas aspek-aspek penting yang harus diperhatikan dalam pembuatan jurnalisme data?

Rahadian : Mulai dari awal sampai terakhir itu mempengaruhi satu sama lain. Jadi dalam setiap tahapan itu punya peran yang sangat penting. Jadi kalau datanya kotor, itu bisa menghasilkan kesimpulan yang juga tidak benar. Jadi ya susah sih. Kalau mau dibilang yang paling penting mana, ya enggak bisa. Menurut saya sih tidak ada yang lebih penting dibanding yang lain. Masalahnya adalah masalah urutan kerja. Mana yang harus lebih dulu dikerjakan. Bukan karena lebih penting tapi ya memang harus itu dulu yang dikerjakan. Data ya memang harus dibersihkan dulu kalau mau dianalisis. Enggak bisa main analisis dulu baru *cleaning* belakangan. Pusing pasti nanti yang bikin analisis. Datanya besar sekali cuma mau analisis sebagian, tapi datanya enggak dipotong dulu, itu problem, istilahnya *knowledge*-nya terlalu tinggi. Gimana kalau enggak dibersihkan dulu, ya enggak akan bisa dapat cerita yang tajam tadi. Karena terlalu tinggi *knowledge*-nya tadi. Karena selalu punya perannya masing-masing yang tahapannya itu harus dijalankan dengan benar. Salah satu enggak benar, pasti ada yang kacau.

Peneliti : Penelitian saya sebenarnya latar belakangnya kan berdasarkan dari banyaknya berita *hoax*, bohong yang beritanya tidak akurat. Sedangkan jurnalisme data ini diidentikkan dengan jurnalisme berita-berita yang lebih akurat, lebih komprehensif, jelas, detail dan lengkap. Lalu bagaimanakah pengaruh jurnalisme data dalam memberantas berita-berita *hoax* seperti itu mas?



Rahadian : Berita data itu kan tidak bisa dianalisis dalam 5 menit. Sementara dalam 5 menit anda bisa menemukan ada beberapa berita *hoax*. Jadi problemnya kalah cepat aja. Karena untuk bikin berita *hoax*, dalam sejam aja ada berapa *hoax* yang bisa diproduksi orang. Sementara yang harus nulis bikin berita untuk data kayak gini enggak bisa sehari dua hari. Jadi akan selalu kalah langkah sih

pada akhirnya, ya *mindset*-nya begitu. Jurnalisme data bukan obat mujarab kalau untuk melawan *hoax*. Dia bisa dipakai sebagai pembanding, tapi karena dia selalu kalah cepat ya itu problem yang tidak bisa dihindari. Karena kalau mau adu cepat, ya enggak akan mungkin. Semua proses ini dijalani dengan supervisi yang benar, itu pasti butuh waktu, itu tren. Untuk menjadikan suatu yang bagus ya butuh waktu. Ketika harus dikerjakan dengan sampah itu ya mau gimana-gimana enggak bisa. Ya solusinya bukan itu. Ini bukan obat mujarab untuk sindrom *hoax*. PR-nya bukan di situ, problemnya bukan di situ. Kalau yang baca percaya *hoax* ya sampai kapan akan jadi masalah. Problemnya kan kita di literasi. Literasinya rendah. Dia enggak bisa membedakan mana yang *hoax* atau bukan. Dia enggak punya level (*scepticism*) yang baik untuk mencerna informasi. Jadi kuncinya bukan, kasarnya saya membuat data *driven*, kami membuat data *driven* bukan untuk *head to head* sama berita *hoax*. Enggak gitu juga. Kita sadar betul, kita enggak saingan sama mereka. Kadang-kadang mereka jauh lebih menarik, lebih *appearing*, lebih emosional. Sementara kita untuk bikin itu menjadi menarik itu jungkir baliknya lama. Tapi paling tidak ada yang bisa dijadikan patokan. Jadi pasti akan ada *delay*. Waktu ini nih yang pas. Dia mungkin bisa membantu mengembangkan, tapi kita yakin persis tidak akan bisa seimbang. Akan selalu kalah, mau seberapa pun yang diproduksi. Karena yang bisa bikin data *driven* siapa, yang bisa bikin *hoax* siapa? Jauh lebih banyak yang bikin *hoax* daripada data *driven*. Mau seluruh media di Indonesia bikin data *driven* sekalian ya akan kalah sama berita *hoax*. Jadi enggak *make sense* kalau dikatakan bisa memberi obat. Kasarnya kalau dalam penyakit itu hanya bisa mengurangi rasa sakit bukan menyembuhkan penyakit. Kalau mau menyembuhkan penyakitnya ya tekor masalahnya. Tekor masalahnya ya bukan urusan media semata-mata. Itu urusan pendidikan. Mangkannya pendidikan yang harus dibenerin kalau tentang *hoax*. Jangan mencoba melawan *hoax* dengan menyebar anti *hoax* lagi. Enggak akan pernah bisa. Mereka yang produksi macem-macam *hoax*. Jadi kita tidak bisa dan menurut saya akan sia-sia kalau kita harus sama mereka. Jadi proporsionalnya adalah ya kalau saya logikakan ini orang sakit ya tadi. Mungkin ini hanya sampai mengurangi rasa sakitnya tetapi virusnya enggak.

- Peneliti : Lalu apakah ada hubungannya antara jurnalisme data dengan berita yang berkualitas? Lalu adakah indikator berita itu berkualitas atau tidak?
- Rahadian : Enggaklah. Kualitas media itu mau pakai ukurannya siapa? Paling gampang kan kode etik jurnalistik. Selama dia memenuhi standart kode etik jurnalistik itu sudah dianggap berkualitas. Serendah-rendahnya, yang penting lolos kode etik itu udah bagus. Kalau mau yang (...) itu tentu lain. Tapi standar minimumnya ya bisa sesuai kode etik. Dalam jurnalistik tidak akan dipakai juga untuk mengklaim itu berkualitas ya enggak juga. Tetap aja harus ada standart yang memenuhi kode etik. Karena yang digunakan sebagai ukuran publik ya jurnalistik. Si jurnalis itu juga harus sepaham kode etiknya gimana. Enggak serta-merta itu jadi kurang, itu apanya yang paling berkualitas. Ini kan hanya *general line* aja.

Lampiran 7.

Surat Permohonan Wawancara dan Penelitian

	UNIVERSITAS MUHAMMADIYAH YOGYAKARTA FAKULTAS ILMU SOSIAL DAN ILMU POLITIK PROGRAM STUDI ILMU KOMUNIKASI Status : AKREDITASI A Nomor : 203/SK/BAN-PT/AK-XVI/S1/X/2013 Alamat : Kampus Terpadu Ngebel, Tamantirto, Kasihan, Bantul Yogyakarta 55001 Telp. (0274) 387656 (hunting)	
No	: 264/KJK/VI/2018	Yogyakarta, 11 Juni 2018
Hal	: Permohonan Wawancara dan Penelitian	
Lampiran	: -	
Kepada Yth.		
Redaksi Beritagar.id		
di Jl. Jatibaru No. 28		
Jakarta Pusat 10160		
Assalamu'alaikum Wr.Wb.		
Sehubungan dengan adanya tugas akhir kuliah (Skripsi), melalui surat ini maka saya selaku mahasiswa Prodi Ilmu Komunikasi Universitas Muhammadiyah Yogyakarta:		
Nama	: Ahmad Galang Marufa	
NIM	: 20140530125	
Email	: marufa.galang@gmail.com	
No. HP	: 082135600123	
Memohon izin untuk melakukan wawancara mengenai Jurnalisme Data dalam Beritagar.id . Adapun data dan informasi yang dicari adalah tentang jurnalisme data dan penerapannya dalam Beritagar.id. Data yang didapatkan kelak hanya untuk keperluan penelitian saja.		
Demikian yang dapat saya sampaikan. Atas izin dan kerja samanya, saya mengucapkan terima kasih.		
Wassalamu'alaikum Wr. Wb.		
		
Mengetahui, Ketua Prodi Ilmu Komunikasi		
Haryadi Arief Nuur Rasyid, S.IP, M.Sc		

Lampiran 8.

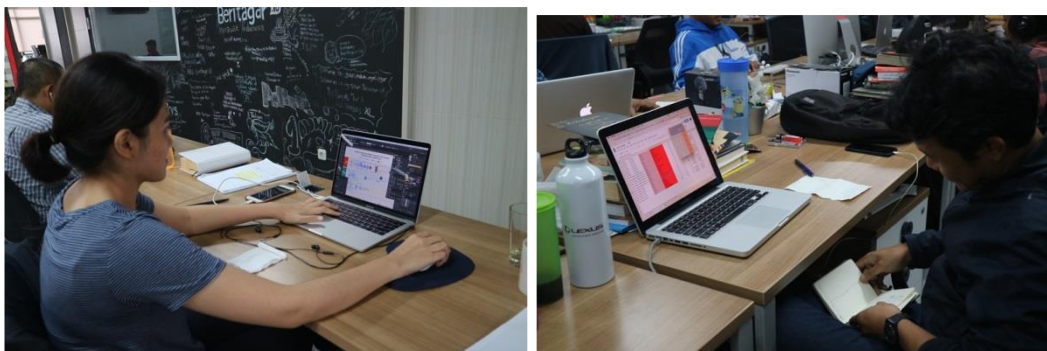
Dokumentasi Foto



Keterangan: Foto peneliti (kiri) bersama dengan Aghnia Adzkia (tengah) dan Rahadian Prajna Paramita (kanan)



Keterangan: Foto suasana ruang redaksi dan ruang kerja di kantor Beritagar.id



Keterangan: Foto suasana redaksi sedang bekerja



Keterangan: Tampak dalam dan tampak luar kantor Beritagar.id yang berlokasi di Jakarta Pusat

Lampiran 9.

Data Mentah Sajian Data Berita dalam Beritagar.id Januari – Maret 2017

Data mentah pada Sajian Data Berita dalam Beritagar.id Periode Januari – Maret 2017 adalah hasil pengumpulan data mentah yang dilakukan peneliti terhadap berita yang dipublikasikan pada Januari – Maret 2017 sebanyak 2.172 berita. Dikarenakan data mentah tersebut berjumlah sebanyak 133 halaman, oleh karena itu data mentah tidak dapat disajikan sepenuhnya dikarenakan terlalu banyaknya halaman. Meski demikian, berikut ini adalah *link* yang dapat diakses untuk melihat data mentah asli:

https://drive.google.com/file/d/12z01o9NT0SzaIWXu3GNT2vOTjf_A5WNI/view?usp=sharing

Pada halaman selanjutnya adalah beberapa halaman sajian data berita dalam Beritagar.id periode Januari – Maret 2017 sebagai contoh.