

Summary

Di era Industri 4.0 budaya digital masyarakat Indonesia sangat dipengaruhi oleh penggunaan media sosial. Dampak dari penggunaan media sosial ini dapat dirasakan masyarakat tidak hanya di dunia maya, namun juga di kehidupan sehari-hari. Diantara topik perbincangan yang membawa dampak paling besar adalah pembahasan isu (topik) agama di media sosial.

Indonesia beberapa tahun belakangan ini merasakan gejolak politik yang jauh lebih tinggi dari tahun-tahun sebelumnya. terlepas dari pro atau kontra (keberpihakan), tidak dipungkiri bahwa isu agama yang bercampur dengan strategi politik di media sosial menjadi katalisator perubahan signifikan suasana perpolitikan dan bahkan kehidupan masyarakat secara umum. Tidak hanya itu berbagai kalangan juga mulai mensinyalir berkembangnya radikalisme di Indonesia juga terkait dengan penggunaan media sosial.

Mengingat besarnya dampak isu agama di media sosial dalam kehidupan bermasyarakat di Indonesia (terutama kaum milenial), maka melalui penelitian ini telah dilakukan kajian mendalam karakteristik perbincangan topik Agama dari waktu ke waktu (*Social Media Analytics*). Penelitian ini menggunakan data percakapan tentang topik agama dari salah satu media sosial populer di Indonesia (twitter) dari tahun penelitian berjalan hingga 10 tahun ke belakang. Total data yang digunakan berjumlah jutaan tweet yang berasal dari ratusan ribu pengguna. Penelitian kuantitatif dengan skala data sebesar dan dalam jangka waktu sepanjang ini adalah yang pertama kali dilakukan di Indonesia atau konteks ke-Indonesiaan.

Kami menganalisa lebih dari 2 juta status media sosial (*posts*) terkait isu agama di twitter selama kurun waktu 2009 hingga 2019 menggunakan berbagai metode seperti text analytics, topic modelling, dan social network analysis. Dari hasil analisa data yang dilakukan lebih nampak bahwa isu agama di politik hanya hadir di tahun-tahun politik. Hal ini mengindikasikan adanya penggunaan isu agama di politik. Terlihat juga pada hasil penelitian bahwa interaksi antar komunitas agama di media sosial masih dipengaruhi dengan kuat oleh efek *echo chamber*. Dimana antar paham keagamaan masih kurang adanya interaksi, hal ini

memperkuat hasil dari penelitian-penelitian sebelumnya dan menambahkan informasi bahwa isu ini semakin terlihat di tahun-tahun politik.

Kata Pengantar

Alhamdulillah atas segala limpahan rahmat dan karuniaNya penulis dapat menyelesaikan penelitian ini sesuai dengan waktu yang telah direncanakan. Sebagai manusia biasa penulis menyadari adanya kekurangan dan kelemahan pada hasil penelitian ini yang masih jauh dari harapan para pembaca. Untuk itu demi perbaikan karya tulis ini dengan segala kerendahan hati penulis menerima segala bentuk kritikan dan saran dari para pembaca sekalian.

Sejak merencanakan proposal sampai dengan terselesainya penelitian ini penulis banyak mendapat dukungan dari berbagai pihak. Untuk itu pada kesempatan yang baik ini penulis menyampaikan ucapan terima kasih yang setinggi-tingginya kepada :

1. Jajaran Pimpinan UIN Syarif Hidayatullah Jakarta: Rektor, Kepala Pusat Penelitian dan Penerbitan (PUSLITPEN) LP2M, Ketua Lembaga Penelitian dan Pengabdian kepada Masyarakat (LP2M) beserta seluruh staff dan jajarannya.
2. Rekan-rekan sesama dosen dan staff yang telah membantu dalam memperoleh informasi yang diperlukan untuk menyelesaikan penelitian ini.

Kiranya hanya Allah SWT yang senantiasa dapat membalas segala kebaikan yang telah diberikan kepada penulis dan penulis berharap semoga karya tulis ini dapat memberikan manfaat bagi perkembangan ilmu Matematika dan Data Science pada umumnya serta kepada penulis pada khususnya.

Jakarta, September 2020

Penulis.

Daftar Isi

Lembar Pengesahan	i
Pernyataan Bebas Plagiasi	iii
Kata Kunci	iv
Summary	v
Kata Pengantar	vii
Daftar Gambar.....	14
I. Pendahuluan.....	15
I.1. Latar Belakang Masalah	15
I.2. Batasan dan Perumusan Masalah.....	16
I.3. Tujuan Penelitian	17
I.4. Manfaat dan Kegunaan Penelitian	17
I.5. Signifikansi dan Novelty	17
II. Kajian Teori.....	19
II.1. Agama di Media Sosial	19
II.2. Analisis Media Sosial	20
II.3. Hipotesis Penelitian	21
III. Metodologi Penelitian	22
III.1. Text Analytics	22
III.2. Topic Modelling.....	23
III.3.1. Evaluasi	27
III.3. Analisis Jaringan Sosial	29
IV. Hasil Penelitian dan Pembahasan	30
IV.1. Data dan Sumber Data	30
IV.2. Studi Empiris	32
IV.2.1. Text Analytics	32
IV.2.2. Topic Modelling.....	36
IV.2.3. Social network Analytics	36
IV.3. Keterbatasan Penelitian.....	37

V. Kesimpulan	39
V.1. Implikasi.....	39
V.2. Rekomendasi	40
VI. Pelaksana Penelitian.....	41
VII. Daftar Pustaka.....	42
VIII. Acknowledgement	44
IX. Lampiran	45
IX.2. Code Yang Digunakan.....	46
IX.3. Draft Publikasi	64

Daftar Gambar

Gambar 1. Contoh text analytics dari berbagai media sosial.	22
Gambar 2. Diagram keterhubungan penelitian terdahulu dengan penelitian yang dilakukan.	24
Gambar 3. Proses kerja <i>probabilistic graphical model</i>	25
Gambar 4. Contoh gambar keterhubungan user di media sosial.	29
Gambar 5. Distribusi data lintas tahun yang digunakan dalam penelitian.	31
Gambar 6. Peta persebaran data <i>Twitter</i> yang digunakan di penelitian ini secara spasial.	32
Gambar 7. Text analytics dari seluruh data.	33
Gambar 8. Frekuensi kumulatif dari penggunaan hashtags dari data yang ada.	34
Gambar 9. Trend isu yang muncul dari tahun ke tahun pada data yang digunakan.	35
Gambar 10. User dengan jumlah follower yang sedikit tetap memiliki potensi besar untuk menghasilkan status dengan engagement yang sangat tinggi.	35
Gambar 11. Topik optimal yang terbentuk dari data yang ada.	36
Gambar 12. Hubungan antar user dan bagaimana efek echo chamber nampak.	37

I. Pendahuluan

I.1. Latar Belakang Masalah

Indonesia adalah salah satu negara yang memiliki banyak keragaman didalamnya seperti suku bangsa, ras, budaya, dan juga agama. Keberagaman ini dipersatukan oleh landasan negara Pancasila yang sangat menjunjung tinggi nilai kebhinekaan sebagai kekayaan. Namun demikian menjaga keseimbangan antara nilai keagamaan dan keberagaman merupakan isu komunal yang masih menjadi salah satu potensi konflik yang cukup tinggi di Indonesia (Ali-Fauzi et al., 2009). Di era informasi, potensi konflik keagamaan ini bertambah dimensinya dengan hadirnya kompleksitas adanya penggunaan teknologi media sosial dan penggunaannya dalam kancah perpolitikan di Indonesia (Kurniawan, 2018).

Di Indonesia pengguna aktif media sosial yang berjumlah lebih dari 160 juta dan presentasi penetrasi hampir 60% (Kemp, 2020). Dengan pengguna yang berjumlah sangat besar ini ditambah dengan sifat media sosial yang menghilangkan jarak antar pengguna, maka media sosial sebenarnya memiliki potensi untuk menjadi jembatan diskusi antar berbagai paham keagamaan untuk saling mengerti dan menghormati. Namun demikian dampak dari kecerdasan buatan (algoritma) di media sosial dan kebebasan tanpa batas justru seringnya membawa dampak negatif yang saling menguatkan yang biasa disebut sebagai *echo chamber* (Neo, 2016).

Untuk mendapatkan gambaran yang lebih jelas serta terukur sejauh mana diskusi tentang agama dan berbagai isu atau topik terkait lainnya di media sosial (online) mempengaruhi dinamika kemasyarakatan di Indonesia, perlu dilakukan kajian empiris analisa media sosial. Namun dengan jumlah pengguna media sosial yang begitu besar analisa data pada permasalahan ini mengalami tantangan dimensi dan keragaman data yang tidak kecil. Permasalahan pengelolaan data yang besar ini biasa disebut sebagai permasalahan Big Data dan tantangan keragaman data (teks) biasanya ditangani dengan ilmu sains data (Ghani et al., 2019).

Meskipun telah ada beberapa studi yang dilakukan terkait internet dan penyebaran pemahaman keagamaan di dunia maya, ada beberapa kelemahan dari studi-studi tersebut. Pertama, studi-studi yang ada cenderung membatasi pada satu tokoh atau satu organisasi saja sehingga sulit untuk memahami kecenderungan umum paham keagamaan yang berkembang di dunia maya. Kedua, secara metodologi belum banyak studi yang menggunakan big data untuk analisis konten dan jaringan sosial dalam penyebaran pemahaman keagamaan di media sosial. Ketiga, belum banyak studi yang mencoba mengeksplorasi secara komprehensif hubungan

antara konteks sosial politik dan perkembangan narasi keagamaan di dunia maya. Oleh karena itu, untuk mengisi kekosongan dalam literatur, penelitian ini bermaksud untuk menganalisa paham dan narasi keagamaan yang berkembang di media sosial, menganalisa jaringan sosial untuk melihat persebaran paham keagamaan online dan menganalisa pengaruh konteks sosial politik bagi perkembangan dan perubahan narasi keagamaan di dunia maya.

Untuk meneliti tentang dinamika topik perbincangan di media sosial metode seperti Topic Modelling dapat digunakan (D. M. Blei, Edu, et al., 2003). Sedangkan keterhubungan antar pengguna dapat di analisa lebih mendalam dengan analisa jaringan sosial (Udanor et al., 2016). Diantara kajian terdahulu terkait perilaku masyarakat dan media sosial dilakukan oleh (Cahyono, 2016). Namun kajian ini bersifat kualitatif dan tidak menganalisa langsung data dari media sosial. Kajian lain terkait agama dan media sosial dilakukan oleh Sari (2017) (Sari Puspita, 2017), pada data media sosial Instagram. Namun sama seperti penelitian tersebut sebelumnya kajian ini tidak menganalisa langsung dengan teknik text mining apapun dan masih bersifat kualitatif.

Berbeda dengan penelitian sebelumnya, dalam kajian yang akan dilakukan di penelitian ini akan digunakan teknik kuantitatif pada data teks bergantung waktu. Salah satu kajian terkait model ini adalah penelitian yang dilakukan oleh Fani, dkk (2016) (Fani et al., 2016). ketika mendiskusikan tentang topik percakapan di berbagai komunitas media sosial twitter secara umum. Namun penelitian ini menggunakan data dengan jumlah besar yang terdistribusi pada jangka waktu yang lama (10 tahun). Melalui penelitian ini diharapkan dinamika perubahan perbincangan topik agama di media sosial dari waktu ke waktu dapat dipahami dengan lebih baik.

1.2. Batasan dan Perumusan Masalah

Sebagaimana paparan permasalahan diatas, maka penelitian ini dibatasi hanya pengelompokan halus (*soft clustering/topic modelling*) data tidak terstruktur dalam bentuk teks. Lebih jauh lagi kami akan mengimplementasikan penelitian ini untuk mengolah data media sosial berbentuk teks dalam skala besar. Adapun beberapa permasalahan yang akan diselidiki dalam penelitian ini adalah:

1. Apakah perbedaan (topik) karakteristik percakapan topik agama masyarakat Indonesia dari waktu ke waktu?
2. Topic modelling apakah yang dapat menghasilkan informasi/insight terbaik dari permasalahan ini?

3. Bagaimana keterhubungan antar pengguna media sosial ketika membicarakan topik agama di media sosial?

1.3. Tujuan Penelitian

Sesuai dengan perumusan masalah pada bab sebelumnya, maka penelitian ini bertujuan untuk:

1. Memahami dinamika perkembangan budaya digital masyarakat Indonesia terkait dengan kehidupan beragama di media sosial.
2. Mendapatkan sebuah teknik analisa Topic Modelling baru (novel) bergantung waktu yang dapat memberikan informasi atau insight yang paling baik.
3. Memahami pola keterhubungan percakapan antar pengguna media sosial agar mendapatkan pemahaman yang baik pola interaksi antar mereka.

1.4. Manfaat dan Kegunaan Penelitian

Walau dampak kehidupan beragama di media sosial sangat besar terhadap kehidupan sehari-hari masyarakat Indonesia, namun hingga saat proposal penelitian ini diajukan belum ada literatur yang meneliti tentang kehidupan agama di media sosial secara quantitative. Lebih jauh lagi aplikasi pemodelan topik di banyak literatur saat ini bersifat stationary (tidak bergantung waktu), padahal dinamika masyarakat dalam menggunakan media sosial sangatlah dinamis atas waktu. Penelitian yang diajukan merupakan salah satu pioneer dalam penerapan pemodelan topik bergantung waktu pada data besar dari media sosial dengan rentang waktu yang memiliki jangkauan yang luas. Dari informasi penelitian kuantitatif yang objektif dan komprehensif berdasarkan data (*data driven*) ini diharapkan menjadi acuan banyak penelitian lain serta para pengambil kebijakan dalam mengambil keputusan atau kesimpulan terkait dinamika kehidupan beragama di media sosial.

1.5. Signifikansi dan Novelty

Penelitian terdahulu terkait proposal penelitian adalah penelitian yang terkait dengan pembahasan isu agama di media sosial dan penelitian terkait topic modelling terutama topic modelling pada data teks bergantung waktu. Diantara kajian terdahulu terkait perilaku masyarakat dan media sosial dilakukan oleh Cahyono (2016) (Cahyono, 2016). Namun kajian ini bersifat kualitatif dan tidak menganalisa langsung data dari media sosial. Kajian lain terkait agama dan media sosial dilakukan oleh Sari (2017) (Sari Puspita, 2017), pada data media sosial

Instagram. Namun sama seperti penelitian tersebut sebelumnya kajian ini tidak menganalisa langsung dengan teknik text mining apapun dan masih bersifat kualitatif yang memiliki bias yang cukup tinggi ke peneliti.

Berbeda dengan penelitian sebelumnya, dalam kajian yang akan dilakukan di penelitian ini akan digunakan teknik kuantitatif pada data teks bergantung waktu. Salah satu kajian terkait model ini adalah penelitian yang dilakukan oleh Fani, dkk (2016) (Fani et al., 2016). ketika mendiskusikan tentang topik percakapan di berbagai komunitas media sosial twitter secara umum.

II. Kajian Teori

II.1. Agama di Media Sosial

Media Sosial menjadi suatu alat yang dapat membantu penciptaan atau pertukaran informasi, ide, minat karier, dan bentuk ekspresi lainnya melalui komunitas dan jaringan virtual (Obar & Wildman, 2015). Media Sosial yang digunakan terus mengalami kenaikan perkembangan dalam kehidupan keseharian masyarakat. Di negara Indonesia, media sosial di Indonesia memiliki pengguna aktif berjumlah lebih dari 150 juta pengguna dengan persentase penetrasi sekitar 60% (Kemp, 2020).

Media Sosial kini telah dijadikan sebagai alat dimana informasi populer berasal untuk mengetahui berbagai kabar/informasi. Bentuk informasi yang diperoleh berupa hiburan, ide, diskusi politik, termasuk didalamnya tentang pembahasan khusus tema agama. Makin meningkatnya penggunaan media sosial sebagai alat komunikasi khusus tema keagamaan pun turut mempengaruhi cara orang melakukan kegiatan keagamaan. Beberapa hasil riset menunjukkan bahwa adanya beragam macam praktik keagamaan baik berupa kegiatan berderma hingga kegiatan-kegiatan virtual makin mudah ditemui, dimana orang merasa yakin dapat memperkuat keimanan dan keyakinan (Laney, 2005).

Ada kemungkinan akibat sifat media sosial yang ekspresif, sehingga hal tersebut bisa digunakan untuk memberi dan menerima dukungan spiritual, mengekspresikan keyakinan seseorang, menginformasikan kepada orang lain terkait praktek/ritual, aktivitas, bahkan memberi pemahaman tentang berbagai kegiatan keagamaan. Media sosial juga menjadi media untuk memperoleh panduan spiritual dan rujukan primer dalam beragama, dan secara bersamaan menawarkan hiburan bagi para penggunanya. Oleh karena itu, media sosial membantu untuk memenuhi kebutuhan dan gairah keagamaan yang bisa diekspresikan dan dipenuhi secara *online* (Brubaker & Haigh, 2017).

Para penganut agama bahkan memiliki keleluasaan dalam memilih pola-pola pengkonsumsian agama secara daring. Mereka bisa memilih berbagai situasi *virtual* sesuai seleranya, baik yang mampu menghadirkan suasana nyaman dalam berdiskusi di komunitas *virtual* religius, yang bisa memunculkan ruang berdoa bersama secara *online* sebagai bentuk kesalehan yang baru, atau bisa saja menjadi alternatif di tengah kejenuhan narasi dakwah yang direproduksi oleh otoritas keagamaan tradisional. Namun demikian ada juga yang memanfaatkan media sosial untuk saling hujat antar masyarakat beragama—yang ternyata

kerap didukung oleh para pemegang otoritas keagamaan bercorak fundamentalis di dunia maya (Van Zoonen et al., 2010).

Media sosial berguna juga untuk mentransmisi paham dan ideologi keagamaan yang terjadi di tengah fenomena keagamaan *online* misalnya, untuk menyebarkan pemahaman konservatif, fundamentalisme beragama, radikalisme, islamisme, hingga ekstremisme (O'Hara & Stevens, 2015). Bahkan pada tingkat yang ekstrim, kemajuan teknologi informasi dan komunikasi ini membantu kelompok terorisme untuk mengembangkan jaringannya dan memobilisasi individu-individu agar melakukan berbagai tindakan teror baik *online* maupun *offline* (Jurriëns & Tapsell, 2017).

II.2. Analisis Media Sosial

Penelitian ini akan menggunakan beberapa analisis data. Data yang berasal dari *Twitter*, analisis secara umum yang digunakan meliputi analisis data tidak terstruktur dan data terstruktur. Pada data tidak terstruktur, analisisnya meliputi beberapa hal antara lain: pertama, *text analytics* atau analisis teks atau dikenal juga dengan istilah *text mining*, merupakan proses otomatis untuk memperoleh informasi penting dari data text yang tidak terstruktur sehingga membentuk pola dan tren yang memiliki arti.

Kedua, *hashtag analytics*, merupakan suatu cara analisis *hashtag* untuk melihat tingkat popularitas suatu *hashtag* keagamaan, hubungan antar *hashtag* keagamaan dan pola persebarannya. Ketiga, terkait dengan *topic modelling*, merupakan jenis analisis pemodelan statistik yang digunakan untuk mengklasifikasikan teks dalam dokumen ke dalam topik tertentu. Keempat, *Social Network Analysis* (SNA) adalah suatu cara analisis struktur sosial yang mengukur hubungan antara individu, kelompok atau organisasi.

Berbeda dengan metodologi ilmu sosial lainnya, menurut perspektif SNA yang menjelaskan bahwa kondisi sosial itu tidak terbentuk dari aktor sosial yang berdiri sendiri. Akan tetapi, terbentuk dari sistem yang terhubung dan saling tergantung yang menciptakan ikatan dan struktur di antara aktor sosial (Knoke & Yang, 2019). SNA turut pula menganalisis posisi dan peran individu dalam jaringan sosial yang diukur dengan memanfaatkan *degree of centrality* - aktor yang memiliki paling banyak hubungan langsung dengan aktor lainnya sehingga menjadikannya aktor sentral dalam struktur jaringan.

Sentralitas aktor ini dihitung dengan menggunakan masing-masing *tweet* keagamaan yang memiliki level *engagement* yang tinggi (total *retweet*, *like*, dan *reply* yang tinggi) atau *centrality by max* dan rata-rata dari semua *tweet* yang memiliki level *engagement* yang tinggi

atau *centrality by mean*. Selain analisis jaringan sosial di level aktor yang memiliki *degree of centrality*, penelitian ini juga melakukan analisis jaringan antara kelompok liberal, moderat, konservatif, dan islamis.

II.3. Hipotesis Penelitian

Berdasarkan kajian pustaka yang telah dipaparkan diatas, maka kami menyusun beberapa hipotesis berikut:

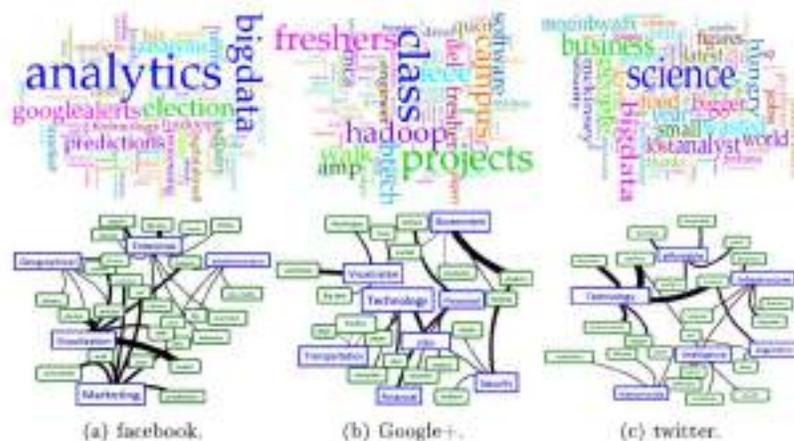
1. Perbincangan tentang topik agama di media sosial mengalami isu komunal dan isu konflik lain yang mengakibatkan polarisasi terutama diakibatkan oleh adanya pengaruh faktor politik praktis di media sosial.
2. Adanya efek echo chamber yang besar di media sosial akan mengakibatkan polarisasi masyarakat di media sosial menjadi semakin meruncing.

III. Metodologi Penelitian

III.1. Text Analytics

Ada beberapa analisis data yang akan digunakan dalam penelitian ini. Untuk data Twitter, secara umum analisis yang digunakan mencakup analisis data tidak terstruktur dan data terstruktur. Analisa data tidak terstruktur dilakukan dengan menganalisa data numerik maupun kategorik yang tersedia di data, misal jumlah like, reply, kapan sebuah status dibuat, dan lain sebagainya. Teknik analisa data terstruktur ini dilakukan melalui serangkaian analisa Exploratory data analysis, yaitu dengan melihat berbagai statistika dasar, visualisasi, dan trend analisis sederhana.

Untuk data tidak terstruktur, analisisnya mencakup beberapa hal antara lain: pertama, text analytics atau analisis teks, merupakan proses otomatis untuk memperoleh informasi penting dari data text yang tidak terstruktur sehingga terbentuk pola dan tren yang bermakna. Metode Text Analytics sebenarnya bagaikan statistika deskriptif atau exploratory data analysis untuk data teks (Hu & Liu, 2012). Text analytics pada penelitian ini sebagian besar dilakukan melalui Voyant Tools (Sinclair & Rockwell, 2020). Adapun tambahan analisa dilakukan dengan melakukan hashtag analyis, yaitu analisis hashtag untuk melihat tingkat popularitas suatu hashtag keagamaan, hubungan antar hashtag keagamaan dan pola persebarannya. Contoh hasil Text analytics dapat dilihat pada Gambar 1 berikut ini.

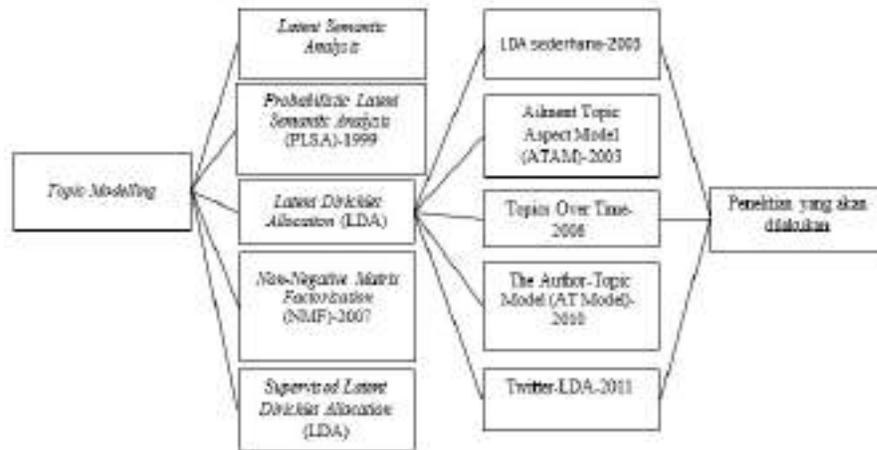


Gambar 1. Contoh text analytics dari berbagai media sosial.

III.2. Topic Modelling

Pemodelan topik adalah algoritma untuk menemukan tema utama yang meliputi koleksi dokumen yang besar dan tidak terstruktur, serta dapat mengatur koleksi dokumen tersebut sesuai dengan tema yang ditemukan (D. M. Blei, 2012). Algoritma ini dapat diadaptasikan ke banyak jenis data seperti data genetik, gambar, dan jaringan sosial (D. M. Blei, 2012). Beberapa bentuk algoritma pemodelan topik seperti pada Gambar 2.2 diantaranya yaitu *Latent Semantic Analysis* (LSA) yang melakukan pendekatan untuk pengindeksan otomatis dan pengambilan informasi yang mencoba untuk mengatasi masalah perbedaan penggunaan kata dengan memetakan dokumen serta istilah untuk representasi (Deerwester et al., 1990)(N. E. Evangelopoulos, 2013)(N. Evangelopoulos et al., 2012), kemudian dilakukan pendekatan baru untuk memperkuat landasan statistik didalamnya dan melakukan analisis faktor yang disebut dengan *Probabilistic Latent Semantic Analysis* (PLSA) (Hofmann & Dan Oneata, 1999)(Hofmann, 2017)(Hofmann, 2001). Dari kedua algoritma tersebut ditemukan suatu permasalahan apabila terdapat matriks pembobotan masing-masing kata memiliki nilai yang negatif, maka dikembangkanlah metode yang disebut dengan *Non-Negative Matrix Factorization* (NMF) (Lee & Seung, 2001)(Li & Zhang, 2007).

Seiring dengan perkembangannya waktu, algoritma pemodelan topik ini terus mengalami perkembangan guna memperoleh hasil pemodelan topik yang baik, dengan keunggulan dalam memperlihatkan topik yang mungkin lebih dari satu pada teks oleh *Latent Dirichlet Allocation* (LDA) membuat salah satu bentuk algoritma pemodelan topik ini banyak diminati didunia penelitian terkait pemodelan topik. Karena menurut para peneliti (D. M. Blei, 2012)(D. M. Blei, Edu, et al., 2003) LDA ini mudah dipahami dan memiliki hasil yang baik dalam pemodelan topik, LDA memiliki banyak pengaplikasian baru contohnya seperti pada Gambar 2 yang dilakukan Michael dan kawan-kawan (Paul & Dredze, 2011), mereka mengembangkan LDA dengan mengkorelasikan pesan *twitter* dengan berbagai aplikasi kesehatan. Pada tahun 2006 (Wang & McCallum, 2006a), LDA kembali dikembangkan untuk data yang memiliki struktur berubah-ubah dari waktu ke waktu. Memasuki tahun 2010, LDA semakin banyak digunakan dalam analisa topik khususnya pada media sosial *twitter* seperti yang dilakukan oleh Hong (Hong & Davison, 2010) yang mengamati kumpulan user dan teks pada *twitter*, kemudian yang dilakukan Wayne Xin Zhao (Zhao et al., 2011) untuk menemukan topik dari *twitter* dengan menggunakan info user.



Gambar 2. Diagram keterhubungan penelitian terdahulu dengan penelitian yang dilakukan.

Pemodelan topik (*Topic Modelling*) adalah algoritma yang digunakan untuk menemukan sebuah tema utama yang terdiri dari dokumen yang besar dan tidak terstruktur, serta dapat mengatur koleksi dokumen tersebut sesuai dengan tema yang ditemukan (D. Blei et al., 2010). Konsep pemodelan topik ini terdiri dari beberapa komponen yaitu kata – kata, dokumen dan *corpora*. Kata dianggap sebagai komponen dasar dari dokumen yang didefinisikan sebagai *item* dari kosa kata yang memiliki indeks untuk setiap kata unik pada dokumen. Dokumen adalah susunan dari kata – kata sebanyak N. Sebuah *corpus* adalah kumpulan dokumen dan *corpora* merupakan bentuk jamak dari *corpus*. Secara sederhana, setiap dokumen dalam corpus mempunyai topik – topik yang dibahas berdasarkan kata – kata yang terkandung di dalamnya (D. M. Blei et al., 2002).

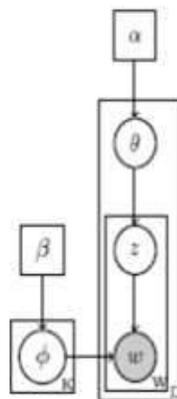
Ide dasar dari suatu pemodelan topik adalah bahwa sebuah topik terdiri dari kata – kata tertentu yang menyusun topik tersebut dan dalam satu dokumen tersebut memiliki kemungkinan terdiri dari beberapa topik yang terkandung dengan probabilitasnya masing – masing (D. Blei et al., 2010). Berdasarkan sudut pandang manusia, dokumen adalah objek yang dapat diamati, sedangkan topik merupakan objek yang tersembunyi sehingga diperlukan suatu metode untuk dapat mengamatinya. Oleh sebab itu, pemodelan topik ini berguna untuk menemukan topik yang tersembunyi pada kumpulan dokumen yang berukuran besar.

Ada beberapa metode yang dapat digunakan dalam melakukan pemodelan topik diantaranya yaitu *Latent Semantic Analysis* (LSA), *Probabilistic Latent Semantic Analysis* (PLSA), *Negative Matrix Factorization* (NMF) dan *Latent Dirichlet Allocation* (LDA). Namun metode yang paling diminati terkait pemodelan topik adalah metode *Latent Dirichlet*

Allocation (LDA) karena dinilai ampuh dalam melakukan analisis topik yang kemungkinan memiliki lebih dari satu topik, serta lebih mudah untuk dipahami (D. Blei et al., 2010). Seiring berkembangnya zaman, LDA pun memiliki banyak pengaplikasian baru. Pada tahun 2006 (Wang & McCallum, 2006b), LDA dikembangkan untuk data yang memiliki struktur yang berubah – ubah dari waktu ke waktu. Memasuki tahun 2010, LDA semakin banyak digunakan dalam analisa topik khususnya pada media sosial *Twitter* seperti yang dilakukan oleh Younghoon Kim (Kim & Shim, 2014) yang menggunakan LDA untuk membuat sebuah sistem rekomendasi pertemanan pada media sosial *Twitter*. Tidak hanya itu saja, LDA juga dapat dimanfaatkan untuk melakukan analisis kecenderungan bisnis dalam dunia perbankan berdasarkan dari kumpulan literatur atau jurnal yang ada di internet (Moro et al., 2015).

Jika masing-masing teks yang sudah dibobotkan sebelumnya mengandung lebih dari satu topik, maka untuk mendapatkan informasi dari teks tersebut dapat dilakukan pengelompokan secara halus atau sering dikenal *soft clustering*. LDA merupakan salah satu contoh metode yang sering digunakan. LDA merupakan suatu teknik mengidentifikasi informasi topik tersembunyi dalam koleksi dokumen besar menggunakan pendekatan *bag of words* (cara representasi data teks) yang memperlakukan setiap dokumen sebagai vektor jumlah kata dan direpresentasikan sebagai distribusi probabilitas atas beberapa topik, sementara setiap topik direpresentasikan sebagai distribusi probabilitas atas sejumlah kata.

Distribusi probabilitas dapat dituliskan dengan struktur bebas bersyarat antar variabel menggunakan grafik agar lebih mudah atau biasa disebut *probabilistic graphical model*. Pada Gambar 3 akan diperlihatkan gambaran proses kerja seperti *probabilistic graphical model*.



Gambar 3. Proses kerja *probabilistic graphical model*

Pada Gambar 3 α dan β merupakan parameter distribusi topik dari dokumen terlebih dahulu atau biasa disebut parameter *prior dirichlet* dan parameter distribusi kata dari sebuah

topik. Nilai keduanya merupakan bilangan riil positif yang dapat dituliskan $0 \leq \alpha, \beta \leq 1$ dan untuk nilai optimalnya didekati secara numerik menggunakan algoritma *Gibbs Sampling* dalam menentukan jumlah topik terdahulu.

Semakin tinggi nilai α mengartikan bahwa setiap dokumen mengandung sebagian besar topik dan semakin rendah nilai α maka dokumen memiliki kemungkinan diwakili oleh beberapa topik. Sementara semakin tinggi nilai β maka suatu topik mengandung campuran sebagian besar kata-kata, sedangkan semakin rendahnya nilai β maka suatu topik hanya mengandung campuran dari beberapa kata.

Seperti pada Gambar 3, distribusi topik dari dokumen (α) mengakibatkan adanya nilai θ sebagai kumpulan campuran topik yang berbentuk matriks probabilitas topik terhadap dokumen seperti pada matriks berikut:

$$A = \begin{bmatrix} \theta_{11} & \dots & \theta_{1K} \\ \vdots & \ddots & \vdots \\ \theta_{M1} & \dots & \theta_{MK} \end{bmatrix}$$

Yang mengartikan bahwa θ_{MK} berarti kumpulan campuran topik ke-K pada dokumen ke-M dan nilai probabilitasnya dapat dihitung menggunakan persamaan berikut:

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{e=1}^K \alpha_e)}{\prod_{e=1}^K \Gamma(\alpha_e)} \prod_{e=1}^K \theta_e^{\alpha_e - 1}$$

Dari kumpulan campuran topik (θ), dapat dipisahkan masing-masing topik (z) dari campuran topik tersebut. Sehingga diperoleh sebuah matriks baru yang berisikan nilai probabilitas kata terhadap dokumen untuk masing-masing topik sebagai berikut:

$$B = \begin{bmatrix} z_{11} & \dots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{M1} & \dots & z_{MN} \end{bmatrix}$$

Dimana nilai z_{MN} mengartikan topik untuk dokumen ke-M pada kata ke-N, dengan nilai $z = \{1, \dots, K\}$ atau dapat diartikan z sebanyak jumlah topik yang ada dari campuran topik, dan nilai masing-masing probabilitasnya dapat dihitung menggunakan persamaan berikut.

$$p(z|\theta) = \prod_{m=1}^M \prod_{e=1}^K \theta_{d,e}^{n_{m,e}}$$

Variabel θ dan z ini biasa disebut dengan parameter laten atau parameter *posterior dirichlet*.

Probabilitas topik yang diperoleh (z) dan distribusi kata pada topik (β) menghasilkan probabilitas kata-kata yang muncul sebagai hasil akhir pembentukan model (w). Sehingga hasil dari model satu dokumen ini akan memunculkan kata-kata dari kelompok yang terbentuk, dan kata-kata ini dapat membantu dalam pendefinisian kategori setiap kelompok. Sehingga total probabilitas berdasarkan grafik model LDA dapat dihitung menggunakan persamaan berikut:

$$p(w, z, \theta | \alpha, \beta) = \prod_{j=1}^M p(\theta_j | \alpha) \prod_{i=1}^k p(\phi | \beta) \prod_{t=1}^N p(z_{jt} | \theta_j) p(w_{jt} | \phi, z_{jt})$$

dengan ϕ merupakan korpus atau untuk kumpulan dokumen yang terjadi akibat adanya β seperti pada Gambar 3. Distribusi peluang bersyarat untuk semua korpus ϕ diformulasikan sebagai berikut:

$$p(\phi | \beta) = \prod_{e=1}^K \frac{\Gamma(\beta_e)}{\prod_{v=1}^V \Gamma(\beta_{e,v})} \prod_{v=1}^V \phi_{e,v}^{\beta_{e,v}-1}$$

dimana V merupakan banyaknya kata. Selengkapnya distribusi peluang korpus w bersyarat z dan ϕ direpresentasikan sebagai berikut:

$$p(w | \phi, z) = \prod_{e=1}^K \prod_{v=1}^V \phi_{e,v}^{n_{e,v}}.$$

III.3.1. Evaluasi

Model yang telah didapat dalam proses LDA harus melalui proses evaluasi untuk melihat tingkat efektivitas model dalam mengklasifikasikan kata. Salah satu teknik evaluasi dalam pemodelan topik adalah dengan melihat nilai dari *coherence score*. *Coherence score* juga dapat digunakan sebagai acuan dalam penentuan jumlah topik yang optimal. Dengan menentukan jumlah topik yang optimal, maka hasil dari klasifikasi tersebut akan semakin baik dalam menginterpretasikan topik.

Coherence score akan mengukur tingkat kesamaan semantik antara kata – kata yang memiliki nilai tertinggi pada topik – topik yang dihasilkan. *Coherence score* memiliki metrik yang berkorelasi dengan interpretasi manusia. Semakin tinggi nilai *coherence score*, maka akan semakin bagus dengan interpretasi manusia (Röder et al., 2015).

Tahapan perhitungan nilai coherence terbagi menjadi empat tahap. Pertama, sekumpulan topik t akan dilakukan segmentasi menjadi sekumpulan kata dari himpunan bagian S . Kemudian probabilitas untuk setiap kata P dihitung berdasarkan data kuantitatif dari *corpus*. Selanjutnya, kedua himpunan P dan S diolah oleh *confirmation measure* untuk menghitung kecocokan dari pasangan S yang menghasilkan nilai ϕ . Terakhir, semua nilai $\vec{\phi} = \{\phi_1, \dots, \phi_{|S|}\}$ dari semua pasangan himpunan bagian S_i di rata – ratakan untuk mendapatkan nilai *coherence*.

Terdapat beberapa teknik yang dapat digunakan untuk menghitung nilai *coherence* seperti C_v , C_{Umass} , C_p , C_{NPMI} , $C_{one-any}$, C_{UCI} dan C_A . Pada penelitian ini, akan digunakan teknik C_{Umass} karena setelah dilakukan beberapa uji coba, teknik ini memiliki kecepatan

pemrosesan yang lebih cepat dibanding teknik yang lain. Probabilitas kata dihitung berdasarkan dari frekuensi dokumen yang digunakan untuk mempelajari topik. Rumus perhitungan *coherence* score menggunakan *Umass* dapat dituliskan sebagai berikut (Röder et al., 2015) :

$$C_{Umass} = \frac{2}{N \cdot (N - 1)} \sum_{i=2}^N \sum_{j=1}^{i-1} \log \frac{P(w_i, w_j) + \epsilon}{P(w_j)}$$

Dimana N merupakan banyaknya dokumen dalam *dataset* dan $w_{i,j}$ merupakan token – token kata unik yang terkandung dalam seluruh dokumen.

Penentuan jumlah topik yang terjadi dalam *topic modeling* merupakan hal yang penting untuk diperhatikan. Dengan penentuan jumlah topik yang sesuai, maka hasil tersebut akan menjadi lebih optimal sehingga mudah untuk diinterpretasi oleh manusia. Untuk menentukan nilai optimal dari banyaknya topik dapat dilakukan dengan menggunakan perhitungan *perplexity*. *Perplexity* dihitung dengan menentukan kemungkinan dari log teks dokumen yang tidak terlihat.

Model yang baik akan memiliki kemungkinan yang tinggi dan menghasilkan nilai *perplexity* yang rendah. Akan tetapi evaluasi menggunakan *perplexity* terkadang tidak berkorelasi dengan dengan baik bagi interpretasi manusia. Cara penentuan banyak topik lainnya ialah dengan menggunakan *topic coherence*. *Topic coherence* menangkap informasi semantik dari topik yang dihasilkan dan menilai interpretasi topik tersebut.

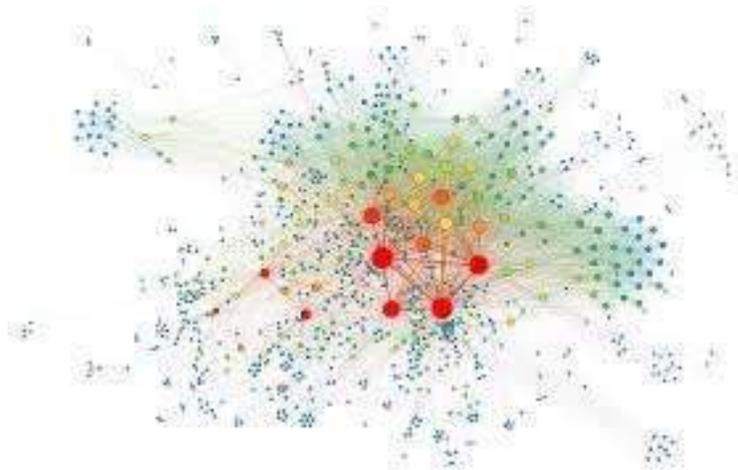
Topic coherence memiliki metrik yang konsisten dengan interpretasi manusia. Semakin tinggi nilai *topic coherence* maka semakin bagus dengan interpretasi manusia. *Topic coherence* dihitung dengan melakukan perbandingan berpasangan antar kata dalam topik tertentu yang mana menghasilkan sebuah ukuran standar kualitas suatu topik

Langkah pertama yang dilakukan adalah sekumpulan topik t dilakukan segmentasi menjadi sekumpulan set pasangan dari subset kata S . Kemudian probabilitas untuk setiap kata P dihitung berdasarkan data kuantitatif dari korpus. Kemudian keduanya P dan S diolah oleh *confirmation measure* untuk menghitung syarat kualitas dari pasangan S menghasilkan ϕ . Terakhir semua nilai $\phi = \{\phi_1, \dots, \phi_{|S|}\}$ di rata-rata kan sehingga menghasilkan nilai koheren.

III.3. Analisis Jaringan Sosial

Analisis Jaringan Sosial (*Social Network Analytics-SNA*) merupakan analisis struktur sosial yang mengukur hubungan antara individu, kelompok atau organisasi. Berbeda dengan metodologi ilmu sosial lainnya, SNA meyakini bahwa kondisi sosial itu tidak terbentuk dari aktor sosial yang berdiri sendiri. Akan tetapi, terbentuk dari sistem yang terhubung dan saling tergantung yang menciptakan ikatan dan struktur di antara aktor sosial (Knoke & Yang, 2020).

SNA membantu menganalisis posisi dan peran individu dalam jaringan sosial yang diukur dengan menggunakan *degree of centrality* - aktor yang memiliki paling banyak hubungan langsung dengan aktor lainnya sehingga menjadikannya aktor sentral dalam struktur jaringan. Sentralitas aktor ini diukur dengan menggunakan masing-masing tweet keagamaan yang memiliki level engagement yang tinggi (total retweet, like, dan reply yang tinggi) atau *centrality by max* dan rata-rata dari semua tweet yang memiliki level engagement yang tinggi atau *centrality by mean*. Selain analisis jaringan sosial di level aktor yang memiliki *degree of centrality*, penelitian ini juga melakukan analisis jaringan antara kelompok liberal, moderat, konservatif, dan islamis. Contoh keterhubungan user di SNA diberikan di Gambar 4. Contoh gambar keterhubungan user di media sosial.



Gambar 4. Contoh gambar keterhubungan user di media sosial.

IV. Hasil Penelitian dan Pembahasan

IV.1. Data dan Sumber Data

Metode kuantitatif digunakan dalam penelitian ini dengan pengumpulan data melalui teknik *crawling* untuk memperoleh data dari platform media sosial *Twitter*. Platform ini dipilih karena ketersediaan data yang terbuka untuk publik merupakan salah satu alasan utama mengapa penelitian ini memilih *Twitter*. Alasan kedua adalah, *Twitter* ditentukan sebagai media sosial yang memiliki penetrasi cukup tinggi dan populer di Indonesia. Laporan dari Datareporter menunjukkan bahwa lima platform media sosial yang terbanyak digunakan Indonesia adalah *Youtube* (88%), *Whatsapp* (84%), *Facebook* (82%), *Instagram* (79%) dan *Twitter* (56%) (Kemp, 2020).

Data dari media sosial merupakan bagian dari *big data* yang memiliki cakupan yang sangat besar. Oleh karena itu, dalam menganalisis perlu adanya pembatasan data. Untuk *Twitter*, kami mengumpulkan data dari tahun 2009 sampai tahun 2019. Untuk pembatasan data, proses sampling yang kami gunakan adalah, sampling berdasarkan pembatasan kata kunci yang kami gunakan untuk proses pencarian data di *Twitter*. Tabel 1 berikut ini adalah beberapa contoh kata kunci yang kami gunakan dalam proses pencarian data di *Twitter*, kata kunci lebih lengkap dapat ditemukan pada lampiran 1 laporan penelitian ini.

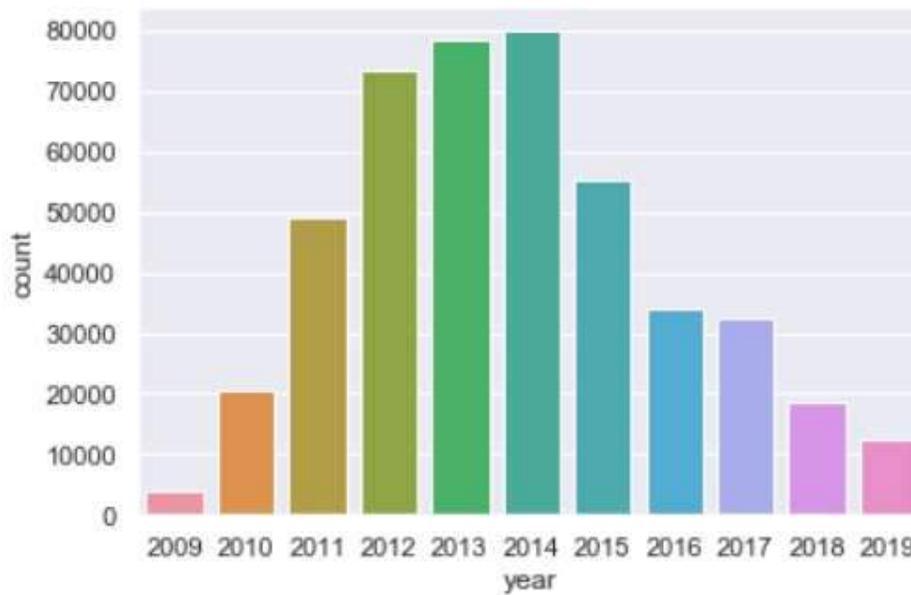
Table 1. Contoh beberapa kata kunci untuk pengambilan data.

allwords	frase	anywords
akidah	azab allah	aqidah
azab	dakwah islam	
berzakat		bersedekah
Bid'ah		bidah
fatwa		ulama, MUI
fikih	aurat perempuan	fiqih
hijrah		
hukum		islam mengenai, islam tentang, islam
hukum		merayakan, mengucapkan,
islam		kaffah, kaffa, politik
Istri		salihah, salehah, solehah, solihah
Jahiliyah		jahil
ketakwaan	apa itu	takwa, ketakwaan, bertaqwa, taqwa

Dengan menggunakan berbagai kata kunci di atas, data yang berhasil kami kumpulkan berupa data *tweet* sebanyak 1.9 juta *tweet*. Akurasi data dipastikan dengan melakukan beberapa tahap *filtering*. Tahap pertama, *filtering* dilakukan secara otomatis menggunakan *machine*

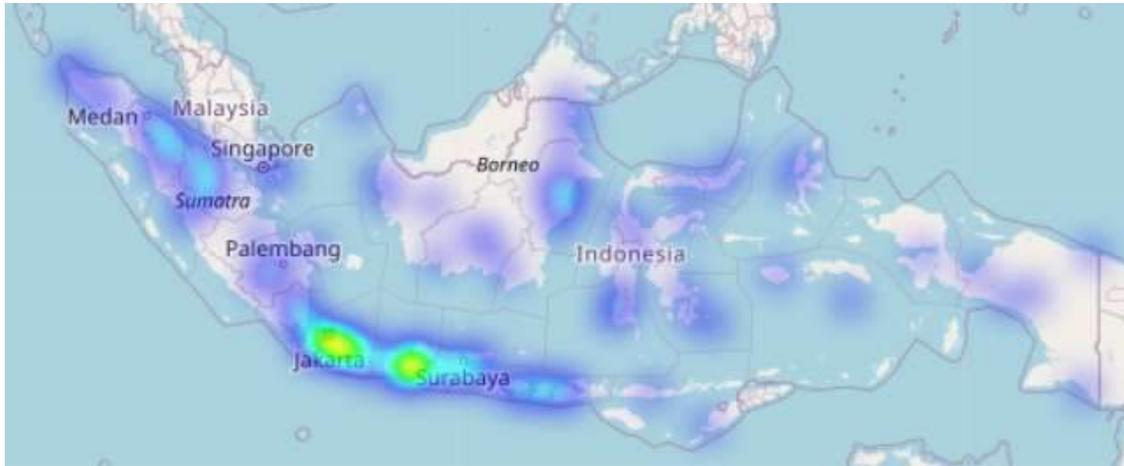
learning dengan beberapa lapisan kode yang berbeda untuk memastikan bahwa hanya *tweet* dari orang Indonesia saja yang akan digunakan dalam analisis. Tahap kedua, *filtering* dilakukan secara manual melibatkan semua tim peneliti untuk memastikan kembali data yang digunakan benar-benar bersih dan akurat.

Setelah proses *filtering*, data yang sudah terfilter dan siap digunakan untuk proses analisis adalah 458,582 *tweet* dari 100,799 *user* (pengguna Twitter), 7,367,190 *follower* dan juga data geospasial pengguna *Twitter*. Distribusi data yang digunakan menunjukkan bahwa pada tahun 2009 tidak terlalu banyak data keagamaan terkait kata kunci pencarian. Sejak tahun 2010, terdapat kenaikan data sampai puncaknya terjadi pada tahun 2014, dan kemudian mengalami penurunan kembali sampai tahun 2019. Detail distribusi data *Twitter* bisa dilihat di Gambar 1 berikut ini.



Gambar 5. Distribusi data lintas tahun yang digunakan dalam penelitian.

Dari sisi lokasi persebaran, data yang diperoleh mencakup seluruh wilayah di Indonesia, meskipun dengan kecenderungan kepadatan (*density*) yang tinggi terkonsentrasi di pulau Jawa. Rincian persebaran data antar wilayah di Indonesia bisa dilihat di Gambar 3 di bawah ini:



Gambar 6. Peta persebaran data *Twitter* yang digunakan di penelitian ini secara spasial.

Meskipun penelitian berbasis data internet (*internet-based research*) memiliki beberapa kelebihan antara lain sumber data yang besar dan kemudahan akses data, jenis penelitian ini juga memiliki beberapa kelemahan. Pertama, penelitian berbasis data internet sangat rentan dengan permasalahan sample bias. Kedua, dibandingkan penelitian *offline*, penelitian berbasis data internet cenderung mengurangi level kontrol peneliti, karena lebih banyak didominasi oleh aspek teknis dan *software* yang digunakan.

Terkait dengan sample bias, pemilihan data media sosial memang membatasi analisis hanya berlaku bagi pengguna media sosial aktif saja. Hal ini tentu saja mengurangi generalisasi dari temuan penelitian ini. Keterbatasan generalisasi dapat dikurangi, oleh karena itu, penelitian ini menggunakan analisis data tidak terstruktur (*unstructured data analysis*) yang bersifat eksploratif dan analisis data terstruktur (*structured data analysis*) untuk memungkinkan adanya inferensi yang dapat digeneralisasi.

IV.2. Studi Empiris

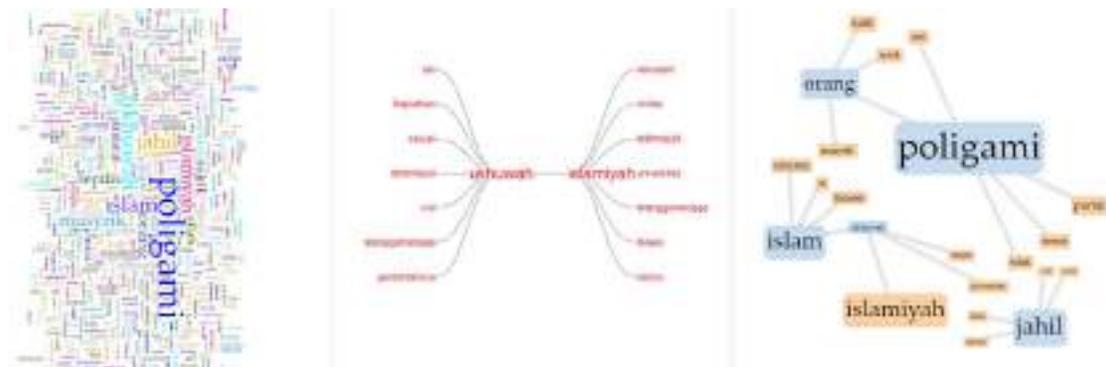
IV.2.1. Text Analytics

Analisa pertama yang dilakukan dalam penelitian ini adalah Text Analytics (TA) menggunakan tools Voyant (Sinclair & Rockwell, 2020). TA pada dasarnya adalah analisa awal statistika deskriptif pada data teks. Sebagai contoh pada Gambar 7 Wordclouds dari keseluruhan data menunjukkan bahwa isu tentang poligami adalah isu yang paling sering dibahas saat pengguna media sosial mendiskusikan tentang agama di twitter di sepanjang kurun

waktu penelitian 2009 hingga 2019. Dari wordclouds juga nampak bahwa isu terkait ukhuwah islamiyah, musyrik dan syirik juga memiliki frekuensi paling tinggi.

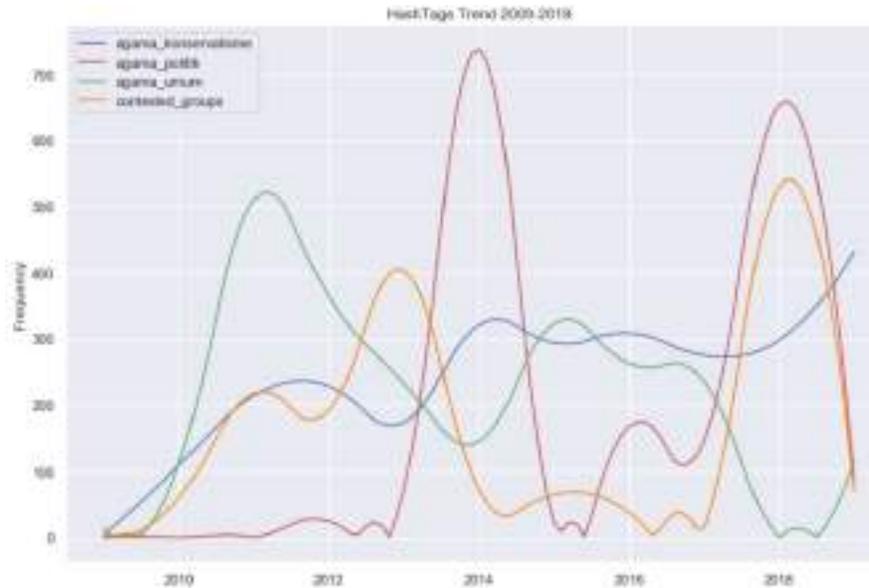
Selanjutnya pada Gambar 7 bagian tengah kita dapat menggunakan WordTree untuk menelaah lebih jauh konteks dari kata-kata yang sering muncul tadi. Dicontohkan pada kasus ini adalah bagaimana topik ukhuwah islamiyah pada temuan wordclouds sebelumnya bisa di jelaskan lebih lanjut ke aspek keindahan ukhuwah islamiyah, pentingnya (kekuatan), dan juga peran universitas dalam menjaga ukhuwah islamiyah. Peneliti menghindari interpretasi terkait ilmu agama dan budaya dan membatasi hanya pada interpretasi dasar dari model yang didapatkan.

Untuk mendapatkan gambaran lebih jelas lagi dari keterkaitan beberapa terminologi atau narasi yang sering muncul di data yang digunakan di penelitian ini, kemudian sebuah model keterkaitan kata (word-links) digunakan (Gambar 7 bagian kanan). Word links menangkap beberapa informasi tambahan yang melengkapi informasi yang sebelumnya di dapatkan di wordclouds dan wordtree. Sebagai contoh isu khilafah ternyata juga muncul secara signifikan pada diskusi tentang ukhuwah islamiyah. Atau adanya faktor politik (partai) pada diskusi terkait poligami. Kemudian dari keterhubungan kata ini yang membentuk tiga cluster yang terpisah menunjukkan ke-3 isu besar ini (poligami, ukhuwah islamiyah, dan jahil) biasanya dibicarakan secara terpisah satu sama lain.



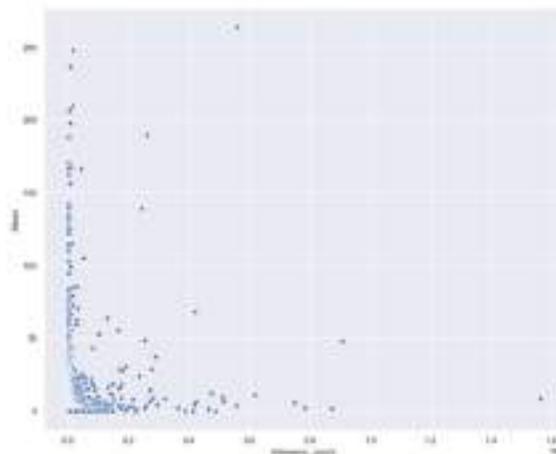
Gambar 7. Text analytics dari seluruh data.

Analisa TA deskriptif yang dilakukan selanjutnya adalah hashtag analysis dari keseluruhan data yang ada. Hashtags adalah kata-kata atau frase kata yang diungkapkan dengan diawali simbol tagar (#). Penggunaan hashtags biasanya ditujukan untuk menunjukkan suatu penekanan tertentu pada sebuah status media sosial. Secara umum hashtags dapat dimanfaatkan dalam analisa data untuk mendapatkan gambaran yang baik tentang topik-topik yang sering dibahas di sebuah media sosial.



Gambar 9. Trend isu yang muncul dari tahun ke tahun pada data yang digunakan.

Meningkatnya konservatisme dan isu politik yang memicu perpecahan yang nampak pada analisa parallel koordinat plot di Gambar 9 merupakan suatu hal yang cukup mengkhawatirkan dan mengancam kesatuan, perdamaian dan keragaman yang ada di Indonesia. Isu ini kemudian di eskalasi lebih jauh lewat hal lain yang ditunjukkan oleh data di Gambar 10. Gambar 10 menunjukkan plot banyaknya follower user di data yang digunakan di penelitian ini pada sumbu horizontal dan rata-rata engagements di sumbu vertikal. Dari Gambar 10 terlihat dengan jelas bahwa begitu banyak pengguna media sosial dengan jumlah follower yang sedikit dapat membuat begitu besar dampak (impact) melalui rata-rata engagements yang begitu tinggi. Hal ini menunjukkan bahwa setiap individu di media sosial dapat menjadi “fire starter” dalam suatu isu besar yang berawal di media sosial.

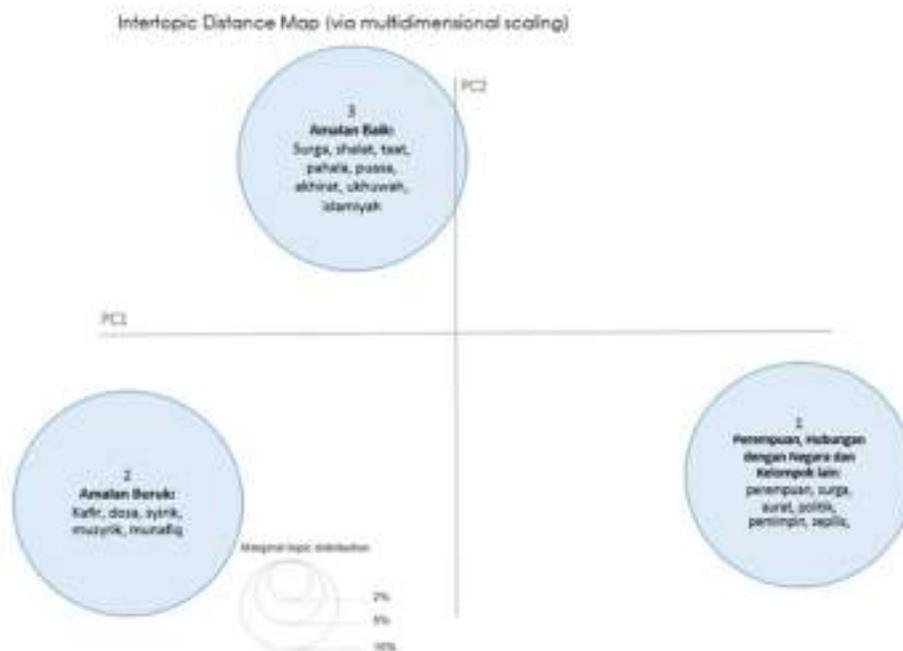


Gambar 10. User dengan jumlah follower yang sedikit tetap memiliki potensi besar untuk menghasilkan status dengan engagement yang sangat tinggi.

IV.2.2. Topic Modelling

Sesuai dengan apa yang telah dijelaskan pada bagian sebelumnya di laporan penelitian ini, peneliti memanfaatkan teknik soft-clustering/topic modelling untuk menangkap topik pembicaraan yang muncul di data penelitian. Metode yang digunakan adalah LDA (D. M. Blei, Ng, et al., 2003) dan menggunakan optimasi parameter nilai coherence antar topik yang terbentuk (Röder et al., 2015). Menggunakan model dan skor evaluasi tersebut didapatkan bahwa jumlah topik yang paling optimal di data adalah sebanyak tiga topik seperti yang nampak pada Gambar 11.

Dari model yang didapatkan nampak bahwa secara umum di data terbentuk tiga buah topik, yaitu topik terkait dengan pembahasan berbagai amalan baik dan buruk, serta topik agama terkait hal lain seperti politik, gender, dan kepemimpinan. Secara statistik maka dapat dikatakan jika terdapat sebuah isu di media sosial, maka besar kemungkinan ia akan termasuk ke dalam salah satu dari 3 isu besar ini.

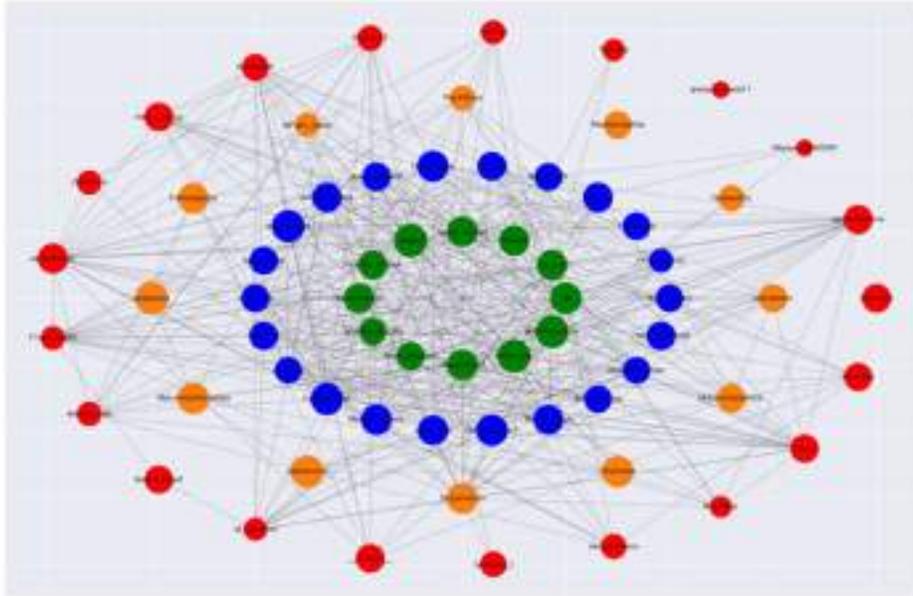


Gambar 11. Topik optimal yang terbentuk dari data yang ada.

IV.2.3. Social network Analytics

Menggunakan data tokoh yang memiliki peran sentral dalam proses konstruksi dan diseminasi narasi keagamaan di media sosial, kami menemukan adanya fragmentasi otoritas

keagamaan di media sosial. Fragmentasi otoritas kami maknai sebagai otoritas untuk menginterpretasi, melegitimasi, mengkonstruksi, dan menyebarkan narasi keagamaan sekaligus membentuk identitas keagamaan tidak hanya dilakukan oleh tokoh agama, tokoh masyarakat, dan organisasi keagamaan resmi, akan tetapi setiap individu dan komunitas yang aktif di media sosial mempunyai peluang dan peran yang sama atau bahkan lebih sentral dari tokoh agama dan masyarakat.



Gambar 12. Hubungan antar user dan bagaimana efek echo chamber nampak.

Gambar 12 menunjukkan adanya fragmentasi otoritas. Lapisan paling dalam (hijau) menunjukkan bahwa akun-akun tersebut adalah akun-akun paling sentral dalam proses dalam proses mengkonstruksi, menginterpretasi dan mendiseminasi narasi keagamaan di media sosial. Dari dalam ke luar, secara berurutan, akun-akun tersebut terdiri dari ustadz (hijau), tokoh masyarakat (biru), komunitas atau akun yang tidak merujuk pada individu tertentu (orange), dan akun yang tidak masuk dalam kategori-kategori sebelumnya (merah). Bahkan, hasil penelitian ini menunjukkan siapa saja bisa memiliki otoritas dan pengaruh yang tinggi meskipun mereka tidak memiliki basis follower yang besar.

IV.3. Keterbatasan Penelitian

Team inti penelitian terdiri dari peneliti dari bidang teknik dan sains. Dalam menginterpretasikan hasil peneliti mendapatkan masukan dari rekan-rekan di bidang ilmu sosial dan keagamaan. Namun demikian kajian lebih mendalam dari hasil yang didapatkan

seharusnya masih dapat dilakukan lebih jauh. Akan tetapi beberapa rencana untuk melakukan kajian mendalam tersebut terkendala adanya pandemi yang terjadi pada saat penelitian ini dilakukan. Di masa depan permasalahan ini juga dapat dihindari dengan mengikutsertakan peneliti dari bidang sosial dan keagamaan dalam team inti.

V. Kesimpulan

Media sosial merupakan media daring (online) yang membawa dampak yang begitu nyata di dunia nyata (luring/offline). Menggunakan sampel diskusi agama di media sosial twitter dengan skala data yang cukup besar (~ 2 juta posts) yang berasal dari ratusan ribu pengguna telah ditunjukkan beberapa indikasi berikut. Diskusi agama umum (murni percakapan agama) di media sosial dari tahun ke tahun trendnya nampak menurun. Penurunan ini semakin jelas saat tahun-tahun politik sedang berlangsung. Pada saat yang bersamaan ini, selain diskusi agama umum menurun, isu komunalitas (konflik) semakin meningkat. Hal ini mengindikasikan adanya penggunaan isu agama dalam politik praktis. Hasil analisis dari penelitian ini juga menunjukkan adanya indikasi bahwa paham konservatis cenderung meningkat dari tahun ke tahun.

Jika dilihat menggunakan topic modelling maka sebenarnya secara umum terbentuk tiga topik utama, yaitu terkait topik perbuatan baik dan buruk, serta topik lainnya yang meliputi isu gender, politik, dan topik lainnya. Dalam penelitian ini juga ditemukan suatu indikasi yang cukup mengejutkan dimana para pengguna media sosial memiliki potensi dampak yang besar walaupun memiliki follower yang sedikit. Isu polarisasi di media sosial juga nampak dari analisa jaringan sosial yang dilakukan. Hal ini selaras dengan hipotesis penelitian dan memperkuat temuan dari beberapa penelitian serupa sebelumnya.

V.1. Implikasi

Penelitian kuantitatif yang menggunakan machine learning dan big data yang dilakukan pada laporan ini termasuk yang pertama kali dilakukan di Indonesia, terutama pada topik terkait agama. Melalui penelitian ini kita bisa mendapatkan gambaran jelas dan terukur dari berbagai fenomena yang sebelumnya sebenarnya sudah dibahas di beberapa literatur, namun belum didukung dengan data yang kuat.

Dari hasil yang disampaikan, terutama terkait berbagai isu konflik akibat politik dan potensi polarisasi (komunal) yang bisa terbentuk dari siapa saja dan bisa berkembang dengan pesat akibat efek echo chamber, maka diharapkan hal ini bisa dimanfaatkan oleh pemangku kebijakan dan masyarakat secara luas untuk mengantisipasi potensi konflik di masyarakat yang bisa berkembang.

V.2. Rekomendasi

Media sosial bagaikan sebuah pedang bermata dua, ia dapat digunakan untuk menyebar kebaikan dan menghilangkan kendala jarak dalam bersosialisasi. Namun kebebasan berekspresi yang nyaris tanpa batas tanpa diiringi dengan edukasi yang baik dapat menimbulkan berbagai masalah seperti yang telah dijelaskan dalam temuan dari penelitian ini. Untuk menangani hal tersebut, peneliti menyarankan sebuah proses yang kontinu untuk menyebarluaskan literasi digital di masyarakat. Begitu pentingnya kompetensi ini, peneliti menyarankan untuk dimasukkan dalam komponen wajib pendidikan formal di sekolah dan atau perguruan tinggi.