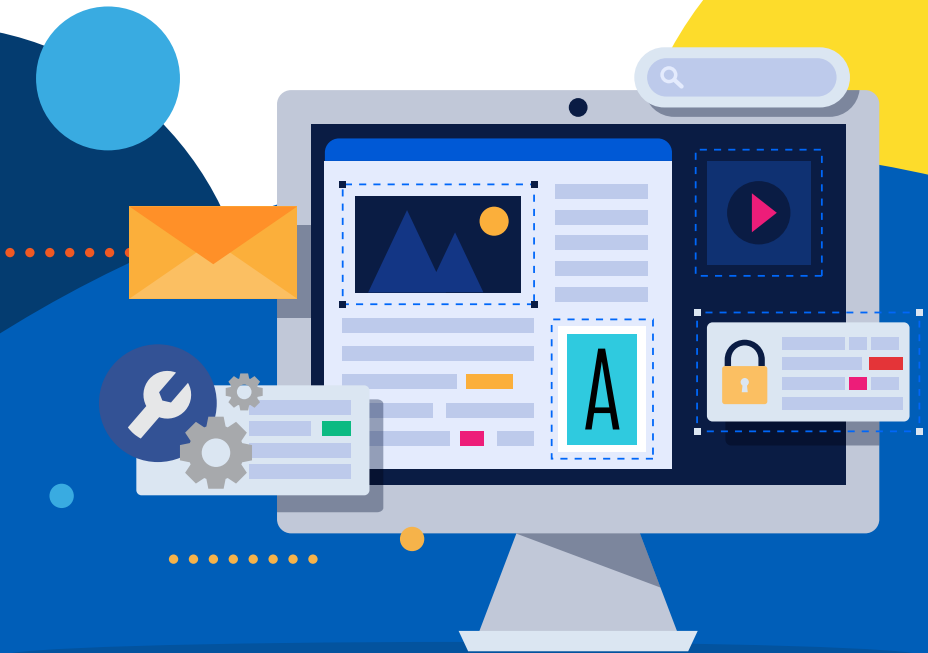
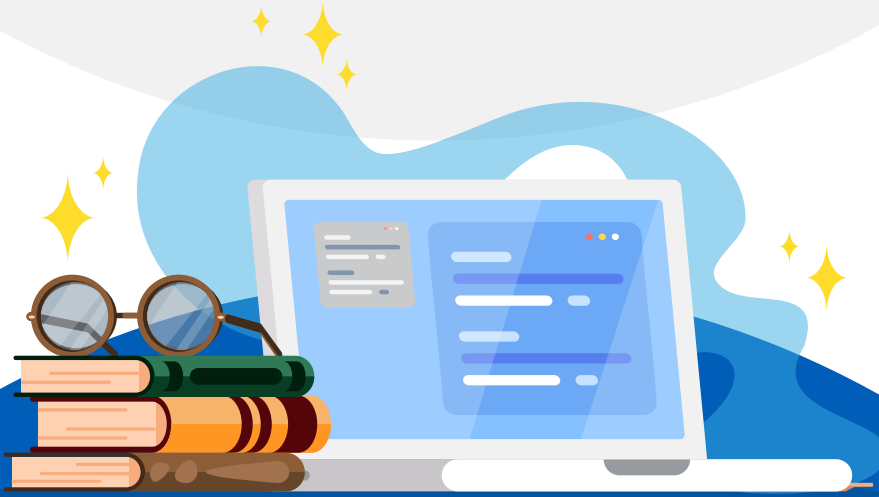


Pengaturan Konten Berbahaya di Indonesia:

Kerangka Hukum, Tren, dan Perhatian





Penulis



Faiz Rahman
Sri Handayani Nasution
Aridiva Firdharizki
Nadya Olga Aletha
Alfredo Putrawidjoyo

Peninjau



Novi Kurnia

Desainer



Riawan Hanif Alifadecya

©2022 Center for Digital Society, Universitas Gadjah Mada.

Hak cipta dilindungi oleh undang-undang. Tidak boleh diproduksi sebagian atau seluruhnya dalam bentuk apapun tanpa izin tertulis dari pengarang dan/atau penerbit.

Penafian

Penamaan yang dilakukan dan penyajian materi di seluruh dokumen ini tidak menyatakan pendapat apapun dari pihak UNESCO mengenai status hukum suatu negara, wilayah, kota, atau daerah mana pun atau otoritasnya, atau mengenai penetapan perbatasan atau batas-batasnya

Ide dan pendapat yang diungkapkan dalam dokumen ini adalah milik penulis; yang belum tentu berada dibawah naungan UNESCO dan tidak terikat terhadap Organisasi.

Dokumen ini dibuat dengan dukungan keuangan dari Uni Eropa. Isinya adalah tanggung jawab penuh penulis dan tidak mencerminkan pandangan Uni Eropa.



Funded by
the European Union



Daftar Isi

I	Daftar Isi
ii	Daftar Tabel
ii	Daftar Gambar
iii	Ringkasan Eksekutif
x	Glosarium
1	Pengantar
2	• Latar Belakang
8	• Metodologi
11	Tinjauan Pustaka
12	• Definisi Konten Ilegal dan Berbahaya berdasarkan Instrumen Hukum Internasional dan Regional
20	• Norma dan Standar Internasional dalam Meregulasi Konten Ilegal dan Berbahaya
28	• Platform dan Moderasi Konten
31	Bab I : Pengaturan Konten Berbahaya Daring di Indonesia
32	• Dasar Konstitusional Pengaturan "Konten Ilegal dan Berbahaya" di Indonesia
35	• Klasifikasi Konten Ilegal dan Berbahaya di dalam Regulasi Indonesia
56	• Metode Penanganan Terhadap Konten Ilegal dan Berbahaya yang Berlaku Berdasarkan Regulasi di Indonesia
60	• Tanggung Jawab Platform Media Sosial dalam Mengatur Konten Ilegal dan Berbahaya

64	Bab II: Tren
65	<ul style="list-style-type: none"> • Kebijakan yang Dikeluarkan dalam Menanggapi Masalah Terkait Konten di Masyarakat
74	<ul style="list-style-type: none"> • Tren Berdasarkan Perspektif Masyarakat
79	<ul style="list-style-type: none"> • Moderasi Konten Berbasis Teknologi dan Otomatis
83	Bab III: Kekhawatiran
84	<ul style="list-style-type: none"> • Mengatur Wilayah Abu-abu: Ujaran Kebencian, Misinformasi dan Disinformasi, serta Pencemaran Nama Baik
87	<ul style="list-style-type: none"> • Kesenjangan antara Mekanisme Regulasi Mandiri Platform dengan Regulasi di Indonesia
93	<ul style="list-style-type: none"> • Dampak Regulasi Konten Terhadap Masyarakat
107	Bab IV : Kesimpulan dan Rekomendasi
108	<ul style="list-style-type: none"> • Kesimpulan
112	<ul style="list-style-type: none"> • Rekomendasi
117	Referensi

→ Daftar Tabel

51	Tabel 1.	Klasifikasi Konten Daring Berdasarkan Peraturan Indonesia
89	Tabel 2.	Perbandingan Klasifikasi Konten dalam Peraturan Indonesia dan Pedoman Komunitas Platform
93	Tabel 3.	Perbandingan Pendekatan Platform Media Sosial dan Negara dalam Menangani Konten Ilegal dan Berbahaya di Indonesia

→ Daftar Gambar

22	Gambar 1.	Tes Ambang Batas Enam Bagian (<i>Six-part thresholds test</i>) dalam Menerapkan Article 20(2) ICCPR
----	------------------	---



Ringkasan Eksekutif

Penelitian ini merupakan bagian dari proyek *Social Media 4 Peace (SM4P)*, yang dilakukan oleh Center for Digital Society, Fakultas Ilmu Sosial dan Ilmu Politik, Universitas Gadjah Mada, bekerja sama dengan United Nations Educational, Scientific and Cultural Organization (UNESCO) dengan dukungan keuangan dari Uni Eropa (UE). Sebagai tahap pertama dari keseluruhan proyek, penelitian ini bertujuan untuk meningkatkan pemahaman tentang kerangka hukum, tren, dan kekhawatiran mengenai regulasi konten berbahaya dan implementasinya di Indonesia untuk lebih memperkuat ketahanan masyarakat dari konten yang berpotensi berbahaya sekaligus memastikan perlindungan kebebasan berekspresi dan berpendapat mereka dalam ruang digital.

Fitur pembagian konten dalam platform internet berbasis *User Generated Content (UGC)* sering disalahgunakan untuk mengirimkan jenis konten yang mungkin berbahaya atau bahkan ilegal. Di Indonesia, isu dis/misinformasi dan ujaran kebencian secara daring—yang menjadi fokus penelitian ini—menjadi semakin mendesak. Pada tahun 2019, disinformasi yang beredar di media sosial selama pemilihan presiden berkontribusi pada kerusuhan yang disertai dengan kasus kekerasan. Kasus ujaran kebencian secara daring terhadap komunitas LGBTQ+ dan minoritas agama juga kerap terjadi. Baik platform media sosial maupun pemerintah telah berusaha untuk meminimalkan kejahatan yang dapat dialami pengguna saat menggunakan internet. Namun, upaya ini bukannya tanpa kekurangan dan/atau dampak terhadap kebebasan berekspresi.

Indonesia telah memiliki beberapa peraturan perundang-undangan untuk konten berbahaya, termasuk mis/disinformasi dan ujaran kebencian, seperti Undang-Undang No. 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik dan revisinya tahun 2016 (UU ITE) dan peraturan pelaksanaannya; Peraturan Pemerintah No. 71 Tahun 2019 tentang

Penyelenggaraan Sistem dan Transaksi Elektronik (PP PSTE); dan Peraturan Menteri Komunikasi dan Informatika No. 5 Tahun 2020 (Permenkominfo 5/20). Namun, undang-undang dan peraturan ini dihadapkan kritik dari organisasi sipil, aktivis, dan akademisi. Beberapa ketentuan di bawah kerangka hukum tentang peraturan konten rentan disalahgunakan oleh pemerintah—penyalahgunaan yang dapat mengarah pada pelanggaran kebebasan berekspresi.

Temuan Utama

→ Kerangka Hukum

Publikasi dan penyebaran konten berbahaya merupakan tindak pidana menurut peraturan Indonesia. Akibatnya, semua konten berbahaya dapat dituntut secara pidana

Tidak ada perbedaan antara konten ilegal dan konten berbahaya di seluruh regulasi mengenai konten di Indonesia. Semua jenis konten yang tercantum dalam peraturan yang ada dipandang secara ilegal dan juga sebagai tindak pidana, termasuk konten yang mengandung disinformasi dan ujaran kebencian. Dengan demikian, hanya ada dua mekanisme yang berlaku dalam menangani konten berbahaya di bawah peraturan Indonesia: mekanisme pengadilan (penuntutan pidana, dan untuk beberapa kasus, tuntutan perdata) dan mekanisme non-pengadilan (penyelesaian sengketa dan tindakan administratif alternatif lainnya).

Beberapa istilah memiliki artian yang terlalu luas (misalnya, moralitas, ketertiban umum, dll.) dan dapat menyebabkan multitafsir, salah tafsir, dan kontroversi

Meskipun banyak kebijakan terkait penanganan konten daring berbahaya telah dikeluarkan oleh lembaga negara dan pemerintah, masih banyak masalah yang muncul karena implementasi UU ITE masih menimbulkan multitafsir dan kontroversi di masyarakat. Hal ini

terlihat dari berbagai kasus terkait konten yang ditangani oleh polisi dan perbedaan interpretasi akan terminologi terkait konten berbahaya dalam banyak putusan pengadilan.

Indonesia mengadopsi pendekatan yang punitif dalam regulasi konten daring

Beberapa regulasi yang mengatur konten juga mengatur platform media sosial, yang secara hukum disebut Penyelenggara Sistem Elektronik (PSE). Peraturan-peraturan tersebut adalah UU ITE, PP PSTE, dan Peraturan Menkominfo 5/20. Peraturan ini menawarkan klasifikasi konten bersama dengan beberapa aturan yang harus dipatuhi oleh platform, termasuk menanggapi permintaan penghapusan konten dalam jangka waktu yang ditentukan. Peraturan tersebut tidak memberikan proses hukum apapun pada masa permintaan take-down, terutama yang dibuat oleh pemerintah. Jangka waktu yang singkat tidak memberikan cukup waktu bagi PSE untuk menilai permintaan penghapusan konten dengan teliti. Pada akhirnya, hal itu dapat memaksa PSE untuk mematuhi agar menghindari hukuman administratif—yang pada prinsipnya menghukum dalam implementasi dan desainnya.

→ Tren dan Kekhawatiran

Disparitas Pengaturan dan Mekanisme Penanganan Konten Berbahaya antara Pemerintah dengan Platform Media Sosial

Di Indonesia, celah ini dimulai dari klasifikasi konten berbahaya. Celah dalam mengklasifikasikan konten menyebabkan perbedaan dalam mekanisme penanganannya. Platform media sosial hanya akan menghapus konten berbahaya yang melanggar pedoman komunitas. Namun, mereka seringkali menggunakan cara lain yang bisa dibilang lebih halus ketika konten tersebut tidak melanggar pedoman, tetapi tetap dapat dianggap berbahaya. Misalnya, platform dapat menandai,

melabel, menurunkan peringkat, dan demonetisasi konten berbahaya. Sebaliknya, Indonesia memandang semua konten berbahaya sebagai konten ilegal. Oleh karena itu, penghapusan konten adalah satu-satunya metode yang diakui secara hukum untuk memoderasi konten 'ilegal' di Indonesia berdasarkan peraturan yang ada.

Pengabaian dan transparansi

Terdapat dugaan bahwa investasi masing-masing platform dalam memoderasi konten dalam bahasa diluar Bahasa Inggris sangat kekurangan dana. Terdapat sangat sedikit informasi mengenai bagaimana platform mempekerjakan moderator konten lokal atau bagaimana mereka melakukan praktik moderasi konten di Indonesia. Di sisi lain, dalam moderasi konten berbasis teknologi atau otomasi yang menggunakan algoritma, mungkin tidak se-efektif yang diklaim oleh platform. Pertama, adanya bias dalam algoritma, yang didapatkan dari bias moderator manusia. Moderasi konten otomatis dapat membuat keputusan yang tidak efektif—bahkan berbahaya—tanpa pertimbangan yang sensitif mengenai perbedaan sosial, budaya, dan politik di seluruh dunia. Dalam praktik mereka di Indonesia, tidak banyak informasi yang tersedia tentang bagaimana platform mengajarkan algoritma untuk moderasi konten berbasis otomasi. Hal ini menunjukkan kurangnya transparansi secara umum dalam praktik platform-platform.

Kurangnya transparansi juga terlihat di pihak pemerintah. Misalnya, informasi tentang permintaan penghapusan konten yang diajukan oleh pemerintah kepada platform tidak tersedia secara publik. Pemerintah juga tidak menyediakan mekanisme banding untuk permintaan mereka.

Peraturan tentang konten secara tidak proporsional merugikan kelompok marjinal

UU ITE dan implementasinya seringkali berpihak pada pelanggaran kebebasan berekspresi, yang jauh dari menjamin perdamaian dan stabilitas bangsa. Undang-undang tersebut dianggap memberikan terlalu banyak kekuasaan dan keleluasaan kepada penegak hukum tanpa proses hukum dan langkah-langkah akuntabilitas yang jelas. Beberapa kelompok masyarakat kerap menjadi sasaran regulasi ini, yakni diantaranya jurnalis dan masyarakat sipil, komunitas Ahmadiyah, minoritas gender dan seksualitas, serta masyarakat biasa. Bagi beberapa kelompok marjinal seperti Ahmadiyah dan komunitas LGBTQ+, peraturan ini dapat melanggar hak-hak mereka, di atas kekerasan struktural dan sistematis yang sudah mereka hadapi setiap hari.

Rekomendasi Utama

Karena masih ada ruang untuk perbaikan moderasi konten, komitmen yang kuat baik dari pemerintah maupun sektor swasta, seperti platform media sosial, sangat dibutuhkan. Beberapa rekomendasi yang diberikan berdasarkan analisis kami adalah:

Rekomendasi 1:

Merevisi UU ITE dan peraturan pelaksanaannya

UU ITE menjadi landasan hukum utama yang mengatur dunia maya di Indonesia. Namun, beberapa hal perlu diperbaiki dalam implementasinya, baik dalam UU ITE maupun peraturan pelaksanaannya. Pemerintah perlu: mempertimbangkan kembali ketentuan klasifikasi konten berbahaya sehingga tidak semuanya dapat dikategorikan sebagai tindak pidana; mendefinisikan kembali istilah yang terkait dengan konten “ilegal” dan “berbahaya”; dan memperbarui mekanisme penanganan konten berbahaya yang ada,

termasuk merumuskan kembali sanksi dan mengembangkan pendekatan komprehensif terhadap konten berbahaya melalui sarana pendidikan dan teknologi.

Rekomendasi 2:

Menyelaraskan hukum dan peraturan terkait konten ilegal dan berbahaya

Keselarasan tersebut diproyeksikan akan mengurangi kemungkinan terjadinya perbedaan interpretasi dan tumpang tindih antar peraturan. Rekomendasi ini tidak hanya dilakukan pada regulasi yang mengatur ranah online, tetapi juga regulasi lain yang mungkin bersinggungan dengan isu konten ilegal dan berbahaya.

Rekomendasi 3:

Menyamakan Persepsi atas Interpretasi Perbuatan yang Dilarang dalam UU ITE

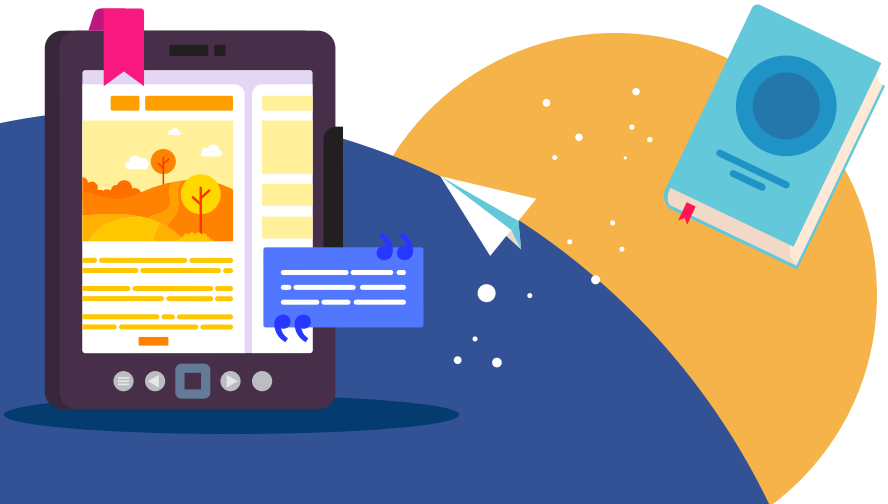
Penelitian ini merekomendasikan Negara untuk memberikan kesamaan persepsi tentang makna ketentuan perbuatan yang dilarang dalam UU ITE. Dalam beberapa kasus, suatu tindakan dapat ditafsirkan berbeda oleh polisi, jaksa, dan pengadilan. Oleh karena itu, harus ada kesatuan persepsi dan interpretasi tentang ketentuan perbuatan yang dilarang. Hal ini dapat dicapai dengan membuat keputusan bersama atau instrumen hukum lainnya mengenai hal ini. Selanjutnya, dalam menentukan apakah konten berbahaya, polisi, jaksa, dan pengadilan juga dapat mengadopsi tes ambang batas enam bagian berdasarkan Rabat Plan of Action, yang termasuk: konteks, pembicara, niat, isi dan bentuk, jangkauan tindakan hasutan, dan tingkat risiko atau kerugian. Selain itu, pelatihan bagi polisi, jaksa, dan pengadilan juga harus dilakukan untuk mendapatkan pemahaman bersama dalam menerapkan ketentuan terkait konten ilegal dan berbahaya.

Rekomendasi 4: Meningkatkan kerja sama antara platform media sosial dan Negara, serta pemangku kepentingan lokal yang relevan dalam menangani konten ilegal dan berbahaya

Negara dan platform media sosial telah berbagi tanggung jawab untuk menangani konten daring berbahaya. Dengan demikian, harus dipastikan bahwa negara dan platform media sosial bergerak ke arah yang sama dalam menangani konten daring ilegal dan berbahaya. Oleh karena itu, berbagai diskusi dan pertemuan antara pemangku kepentingan perlu didorong.

**Rekomendasi 5:
Meningkatkan transparansi dalam memoderasi konten**

Baik negara maupun platform media sosial perlu memastikan transparansi yang menyeluruh dalam penerapan regulasi konten mereka. Transparansi yang menyeluruh berarti bahwa mereka tidak berhenti pada pembuatan laporan berbasis *output* (misalnya, jumlah konten yang dimoderasi,) tetapi juga menginformasikan kepada publik tentang keseluruhan proses moderasi konten.





Glossary

AI	<i>Artificial Intelligence</i> (Kecerdasan Buatan)
AMRI	<i>ASEAN Ministers Responsible for Information</i> (Menteri ASEAN yang Bertanggungjawab untuk Informasi)
APC	Association for Progressive Communications
ARED	<i>Abolition of Race and Ethnic Discrimination</i> (Penghapusan Diskriminasi Ras dan Etnis)
ASEAN	Association of Southeast Asian Nations
CSIS	Center for Strategic and International Studies
COVID-19	Penyakit Coronavirus 2019
DE	Dokumen Elektronik
DSA	Digital Services Act
DUHAM	Deklarasi Universal Hak Asasi Manusia
HRC	<i>Human Rights Committee</i> (Komite Hak Asasi Manusia PBB)
ICCPR	<i>the International Covenant on Civil and Political Rights</i> (Kovenan Internasional tentang Hak Sipil dan Hak Politik)
ICERD	<i>the International Convention on the Elimination of All Forms of Racial Discrimination</i> (Konvensi Internasional tentang Penghapusan Segala Bentuk Diskriminasi Rasial)
ICESCR	<i>the International Covenant on Economic, Social and Cultural Rights</i> (Kovenan Internasional tentang Hak Ekonomi, Sosial dan Budaya)
IE	Informasi Elektronik
IHRL	<i>International Human Rights Law</i> (Hukum Hak Asasi Manusia Internasional)

ITE	Informasi dan Transaksi Elektronik
Kabareskrim Dirtipidsiber	Kepala Badan Reserse Kriminal Direktorat Tindak Pidana Siber
Kemenkominfo	Kementerian Komunikasi dan Informatika
KPU	Komisi Pemilihan Umum
LGBTQ+	<i>Lesbian, Gay, Bisexual, Transgender, Intersex, Queer/Questioning, Asexual and other sexual preferences, and gender identities</i> (lesbian, gay, biseksual, transgender, queer, dan lainnya)
MK	Mahkamah Konstitusi
OHCHR	Office of the High Commissioner for Human Rights
PBB	Perserikatan Bangsa-Bangsa
PDE	Pertukaran Data Elektronik
PN	Pengadilan Negeri
PP	Peraturan Pemerintah
PSE	Penyelenggara Sistem Elektronik
PSTE	Penyelenggaraan Sistem dan Transaksi Elektronik
SAFENet	Southeast Asia Freedom of Expression Network
SM4P	Social Media 4 Peace
TIK	Teknologi Informasi dan Komunikasi
UE	Uni Eropa
UGC	<i>User Generated Content</i>
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
USAID	United States Agency for International Development
UUD 1945	Undang-Undang Dasar Negara Republik Indonesia Tahun 1945
WEF	World Economic Forum



Pengantar

→ Latar Belakang

Pada tahun 2021, jumlah pengguna media sosial mencapai 4,62 miliar secara global, mengalami pertumbuhan 10,1% lebih banyak dari tahun sebelumnya (Hootsuite, 2022). Di Indonesia, sampai Februari 2022, jumlah pengguna aktif media sosial melebihi 191,4 juta, atau sekitar 68,9% dari populasi (Hootsuite, 2022). Jumlah tersebut meningkat sekitar 12,6% dibandingkan tahun lalu. Platform media sosial yang paling banyak digunakan di Indonesia adalah WhatsApp, Instagram, dan Facebook, dengan persentase pengguna internet bulanan sebesar 88,7% untuk WhatsApp, 84,8% untuk Instagram, dan 81,3% untuk Facebook (Hootsuite, 2022). Dengan meningkatnya jumlah pengguna media sosial di Indonesia, potensi pengguna untuk terpapar konten ilegal dan berbahaya juga semakin meningkat.

Dalam platform media sosial, pengguna dapat memproduksi dan membagi konten dalam bentuk gambar, video, komentar, produk, iklan, dan jenis konten lainnya. Konten-konten tersebut biasanya dikenal dengan *User Generated Content* (UGC) (Setiawan, 2021). Fitur *content-sharing* atau berbagi konten ini memudahkan pengguna untuk mengekspresikan ide, bertukar informasi, dan berinteraksi. Sayangnya, fitur ini juga sering digunakan untuk menyebarkan berbagai jenis konten yang mungkin berbahaya atau bahkan ilegal, seperti konten yang mengandung terorisme, eksploitasi seksual, ujaran kebencian, dan misinformasi atau disinformasi.

Pada tahun 2019, Kementerian Komunikasi dan Informatika (Kemenkominfo) Indonesia menerima lebih dari 430 ribu laporan mengenai konten berbahaya, dengan peringkat teratas ditempati konten pornografi, disusul fitnah, konten mengganggu, perjudian, penipuan, mis/disinformasi (hoax), dan lain-lain (Kemenkominfo, 2020). Di antara konten berbahaya yang beragam, persebaran mis/disinformasi dan konten ujaran kebencian di Indonesia— isu yang menjadi fokus penelitian ini—menjadi semakin mendesak di Indonesia.

Penelitian ini menguraikan pemahaman tentang misinformasi dan disinformasi sebagai berikut: misinformasi dan disinformasi keduanya diklasifikasikan sebagai tindakan menyebarkan informasi palsu, meskipun dengan maksud yang berbeda. Untuk misinformasi, orang yang menyebarkannya biasanya tidak tahu bahwa informasi tersebut tidak benar, sedangkan untuk disinformasi, orang tersebut dengan sengaja menyebarkan informasi palsu dengan niat buruk untuk menyebabkan orang lain mendapat informasi yang salah (UNESCO, 2018). Ujaran kebencian adalah tindakan ujaran publik yang berniat buruk yang ditujukan kepada anggota kelompok marjinal secara sistematis terpinggirkan untuk menimbulkan rasa rendah diri, membenarkan perilaku diskriminatif, dan merampas kekuatan kelompok sasaran (Gelber, 2019 diambil dari Sinpeng et al., 2021). Selain itu, penelitian ini juga memasukkan pembahasan pencemaran nama baik di Indonesia dalam laporan ini. Meskipun pemahaman tentang pencemaran nama baik masih samar, laporan ini menemukan urgensi untuk membahas bagaimana pemerintah mengatur jenis konten yang mereka pahami sebagai 'pencemaran nama baik'. Penelitian ini melihat upaya untuk mengatur masalah ini dapat memudahkan lebih banyak pembatasan kebebasan berpendapat di Indonesia.

Kasus mis/disinformasi sudah menjadi perhatian Kemkominfo sejak beberapa tahun lalu. Badan pemerintah secara teratur memperbarui kasus mis/disinformasi melalui situs web mereka.¹ Meningkatnya kasus mis/disinformasi dapat dilihat, misalnya, saat masa pemilu (Katadata, 2020). Selama Pemilu di tahun 2019, mis/disinformasi yang terkait masalah etnis dan agama turut memicu kerusuhan (Temby, 2019).

Konten yang mengandung ujaran kebencian juga menunjukkan peningkatan yang signifikan. Sejak 2018, Kemkominfo menangani dan memblokir akses ke 3640 konten ujaran kebencian daring berdasarkan

¹ Lihat informasi publik Kemkominfo mengenai kasus mis/disinformasi melalui https://epid.kominfo.go.id/informasi_publik/Informasi%20Publik%20Setiap%20Saat.

suku, agama, ras, dan antargolongan (SARA) (Kemkominfo, 2021). Menurut Center for Strategic and International Studies (CSIS), fenomena ujaran kebencian di Indonesia mengalami peningkatan pesat selama satu dekade terakhir (CSIS, 2021). Secara khusus, kasus ujaran kebencian daring terhadap komunitas LGBTQ+ dan minoritas agama dilaporkan lebih sering terjadi (Sinpeng et al., 2021; Lina et al., 2021). Berdasarkan laporan CSIS, jumlah konten ujaran kebencian terhadap komunitas rentan seperti Ahmadiyah, Syi'ah, dan kaum Tionghoa meningkat secara signifikan pada tahun 2021.²

Di sisi lain, selain kasus mis/disinformasi dan ujaran kebencian, jenis konten lainnya yang disebut sebagai konten pencemaran nama baik menduduki peringkat tertinggi kasus terlapor pada tahun 2020 dengan 1743 dari 4656 laporan (CNN, 2020). Laporan dari SAFEnet menunjukkan temuan serupa dan menyatakan bahwa jumlah hukuman terkait konten online berbahaya telah meningkat empat kali lipat dibandingkan tahun sebelumnya, terutama pada pencemaran nama baik, ujaran kebencian, dan mis/disinformasi (SAFEnet, 2021). Namun demikian, kasus mis/disinformasi, ujaran kebencian, dan pencemaran nama baik saling terkait dalam banyak kasus, karena banyak kasus ujaran kebencian atau pencemaran nama baik disebabkan oleh konten mis/disinformasi.³

Indonesia sudah memiliki beberapa undang-undang dan peraturan mengenai konten ilegal dan berbahaya, termasuk, Undang-Undang No. 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE) dan peraturan pelaksanaannya, seperti Peraturan Pemerintah No. 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik (PP PSTE). Namun, undang-undang dan peraturan ini menghadapi kritik dari

² Lihat laporan CSIS mengenai tren kasus ujaran kebencian terhadap komunitas rentan melalui <https://dashboard.csis.or.id/hatespeech/#trends>

³ Sebagai contoh, Lucky Alamsyah dilaporkan ke polisi oleh Roy Suryo tahun 2021 dengan tuduhan pencemaran nama baik karena Alamsyah dituduh menyebarkan berita palsu mengenai Roy Suryo. Lihat CNN Indonesia "Roy Suryo Polisikan Lucky Alamsyah Kasus Pencemaran Nama Baik" <https://www.cnnindonesia.com/nasional/20210524175343-12-646388/roy-suryo-polisikan-lucky-alamsyah-kasus-pencemaran-nama-baik>.

organisasi sipil, aktivis, dan akademisi. Sejak ditetapkannya UU ITE pada tahun 2008, berbagai organisasi, aktivis, dan akademisi mengkritik penerapan UU ITE karena dianggap mengancam kebebasan berbicara (SAFENet, 2021; Amnesty International Indonesia et al., 2021). Sebagian besar kritik terhadap UU ITE berkisar pada proses pembuatannya dan interpretasi terhadap beberapa terminologi “perbuatan terlarang”, termasuk pencemaran nama baik, misinformasi dalam transaksi elektronik, ujaran kebencian, dan penyadapan. Selain itu, sejak tahun 2008, asas konstitusionalitas dari ketentuan tersebut telah beberapa kali diuji melalui pengujian undang-undang ke Mahkamah Konstitusi.⁴ Menariknya, dalam semua putusan yang mengkaji “perbuatan terlarang”, hanya satu yang dikabulkan oleh Mahkamah Konstitusi mengenai intersepsi ilegal. Namun, dalam semua keputusan terkait konten ilegal dan berbahaya seperti pencemaran nama baik dan ujaran kebencian, pengadilan berpendapat bahwa ketentuan ini konstitusional, meskipun kenyataan di lapangan menunjukkan sebaliknya.

Faktanya, kebebasan berbicara dan berekspresi yang dilindungi oleh The International Covenant on Civil and Political Rights (ICCPR) telah dijamin dalam Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (UUD 1945) dan Undang-Undang Hak Asasi Manusia Indonesia (UU No. 39 Tahun 1999). UUD 1945 pasca amandemen memberikan pasal khusus tentang hak asasi manusia, yang juga mencakup perlindungan kebebasan berbicara dan berekspresi. Hal ini juga ditegaskan dalam UU Hak Asasi Manusia (UU No. 39 Tahun 1999) melalui beberapa pasal, yaitu Pasal 14, Pasal 23 ayat (2), dan Pasal 25. Selain itu, UUD 1945 Pasal 28J ayat (2) juga menegaskan bahwa pembatasan hak tersebut harus didasarkan pada undang-undang dan mempertimbangkan tuntutan yang adil berdasarkan moralitas, nilai-nilai agama, keamanan, dan ketertiban umum dalam masyarakat yang demokratis.

⁴ Lihat e.g., Keputusan Mahkamah Konstitusi No. 50/PUU-VI/2008 (kasus pencemaran nama baik), Keputusan Mahkamah Konstitusi No. 2/PUU-VII/2009 (kasus pencemaran nama baik), Keputusan Mahkamah Konstitusi No. 52/PUU-XI/2013 (kasus ujaran kebencian), Keputusan Mahkamah Konstitusi No. 76/PUU-XV/2017 (kasus ujaran kebencian).

Terlepas dari adanya dasar konstitusional dan hukum untuk kebebasan berpendapat serta berekspresi yang signifikan di Indonesia, praktik implementasinya tetap terjadi masalah, terutama di era digital ini. Berdasarkan kerangka hukum yang menjadi dasar untuk mengatur konten, seperti UU ITE, KUHP, dll., definisi (misalnya moralitas, ketertiban umum, dll.) yang membatasi kebebasan berekspresi masih tidak jelas dan berpotensi disalahartikan. Hal ini dapat menyebabkan terjadinya penyalahgunaan yang dilakukan oleh aparat penegak hukum selama masa investigasi kriminal atau proses pengadilan. Selain itu, hal ini juga dapat menimbulkan penyalahgunaan yang dilakukan oleh pemerintah dan/atau otoritas terkait, seperti membungkam kritik terhadap kebijakan pemerintah untuk kepentingan politik, seperti yang sudah disebutkan sebelumnya.

Terlebih lagi, di negara demokrasi, menghormati hak satu sama lain dan ketertiban umum menjadi isu penting dalam pengimplementasian hak atas kebebasan berpendapat (Anindyajati, 2021). Dikarenakan pengakuan global dari kebebasan berekspresi, penting untuk menyelidiki instrumen hukum internasional yang relevan—termasuk standar internasional mengenai konten ilegal dan berbahaya, yang berkisar seputar isu kebebasan berekspresi, terutama di ruang digital.

Sebagai bagian dari pihak yang terdampak dari regulasi konten ilegal dan berbahaya di ruang digital, penting untuk mengungkap kebijakan konten setiap platform media sosial. Namun, terdapat ketidaksesuaian antara peraturan yang ada dengan pedoman platform media sosial (Haryanto, 2020). Salah satu contoh adalah definisi kedua belah pihak mengenai "konten terlarang". "Konten terlarang" tidak disebutkan secara eksplisit melalui UU ITE, tetapi dapat ditafsirkan sebagai bagian dari "Tindakan Terlarang" yang diatur dalam Bab VII. "Tindakan Terlarang" yang terkait konten daring contohnya konten yang mengandung perjudian, pelanggaran moral, penistaan dan/atau pencemaran nama baik,

pemerasan, dan/atau ancaman, serta ujaran kebencian dan berita bohong dalam transaksi elektronik.

Di sisi lain, platform media sosial mendefinisikan "konten terlarang" sebagai "konten berbahaya" dan diklasifikasikan ke dalam konten yang harus dihapus serta dibatasi. Walaupun konten yang harus dihapus sama definisinya dengan pemerintah Indonesia, hal ini berbeda di dalam platform media sosial. Dalam kasus berita palsu, media sosial hanya bisa membatasi penyebarannya dan tidak bisa menghapus sepenuhnya konten dari platform karena sulitnya membedakan berita palsu dan opini.

Untuk meringkas, penting untuk merumuskan kembali regulasi konten digital Indonesia yang ada saat ini untuk menjamin kebebasan berpendapat dan berekspresi. Selanjutnya, perlindungan yang tegas dan kuat terhadap kebebasan berekspresi juga merupakan pondasi penting bagi demokrasi, peraturan hukum, perdamaian, stabilitas, pembangunan yang berkelanjutan dan inklusif, dan partisipasi tinggi dalam urusan publik (Dewan Uni Eropa, 2014).

Penelitian ini merupakan fase pertama dari proyek Social Media 4 Peace (SM4P) secara keseluruhan, yang diselenggarakan oleh Center for Digital Society UGM, bekerja sama dengan UNESCO dan didanai oleh UE. Tujuan dari penelitian ini adalah untuk memetakan hukum, peraturan, dan kebijakan nasional terkait konten ilegal dan berbahaya serta mengetahui penyebab terjadinya masalah yang timbul dari regulasi konten dalam kerangka hukum Indonesia. Penelitian ini juga akan menganalisis kerangka hukum internasional dan kebijakan platform media sosial saat ini terkait dengan konten ilegal dan berbahaya.

➔ Metodologi

Penelitian ini dilakukan dengan melakukan tinjauan pustaka serta pemetaan dan analisis regulasi untuk memahami bagaimana kerangka hukum yang ada mengatur konten daring di Indonesia, bagaimana regulasi tersebut diterapkan, apa dampaknya terhadap kebebasan berpendapat di Indonesia, dan bagaimana hal tersebut mempengaruhi komunitas rentan, seperti komunitas LGBTQ+ dan kaum minoritas agama dalam ranah daring.

➔ Tinjauan Pustaka

Tinjauan pustaka bertindak sebagai dasar untuk penelitian ini; oleh karena itu, hal ini dilakukan untuk memahami beberapa poin penting:

● **Terminologi Konten Ilegal dan Berbahaya**

- Meneliti perkembangan definisi dari konten ilegal dan berbahaya serta norma-norma yang relevan di tingkat internasional dan regional.
- Menentukan penggunaan klasifikasi konten yang 'berbahaya' dan 'ilegal' pada moderasi konten dan mendiskusikan apakah perbedaan ketentuan tersebut memengaruhi penerapan peraturan dan mekanisme moderasi konten.

Norma dan Praktik yang ada mengenai Moderasi Konten

Memahami praktik terbaik mengenai moderasi konten yang diakui secara internasional dan norma-norma yang ditetapkan oleh berbagai organisasi internasional atau regional.

Platform dan Moderasi Konten

Mempertanyakan peran sektor swasta dalam memberhentikan laju persebaran konten ilegal dan berbahaya dalam platform mereka dan meninjau kinerja dan transparansi dalam mekanisme pengaturan mandiri yang diterapkan.

→ Pemetaan Regulasi dan Analisis Regulasi Indonesia Terkait “Konten Berbahaya” dan Implementasinya

Pemetaan regulasi terdiri dari dua analisis, yaitu pemetaan regulasi dan kebijakan serta analisis implementasinya melalui studi kasus. Pemetaan regulasi dan kebijakan dilakukan untuk menggalang beberapa poin penting:

- bagaimana pemerintah mendefinisikan dan mencirikan konten ilegal atau berbahaya;
- solusi yang berlaku dan metode penanganan konten berbahaya;
- tanggung jawab platform media sosial;
- mekanisme pengaturan diri (self-regulatory) yang diprakarsai oleh platform media sosial.

Pemetaan dan analisis regulasi dilakukan dalam empat tahap:

• **Identifikasi serta analisis peraturan utama dan yang terkait mengenai konten daring yang ilegal dan berbahaya.**

UU ITE dan peraturan pelaksanaannya, karena mereka yang terutama digunakan untuk menangani aktivitas online, termasuk konten ilegal dan berbahaya, akan dianalisis. Selanjutnya, peraturan relevan yang ada yang dapat digunakan dalam menentukan dan menangani konten online berbahaya diperiksa berdasarkan jenis konten ilegal dan berbahaya yang disebutkan dalam UU ITE dan peraturan pelaksanaannya.

• **Identifikasi dan analisis keputusan pengadilan terkait konten daring ilegal dan berbahaya.**

Berdasarkan peraturan yang teridentifikasi, putusan pengadilan (baik di Mahkamah Konstitusi maupun Mahkamah Agung) akan diperiksa, untuk mendukung analisis peraturan, terutama bagaimana peradilan menafsirkan dan menerapkan peraturan melalui proses pengadilan.

Mengidentifikasi dan menganalisis kebijakan terkait yang dikeluarkan oleh pemerintah dan lembaga-lembaga negara.

Selanjutnya, melalui analisis regulasi yang ada saat ini, kebijakan yang dikeluarkan oleh pemerintah atau lembaga negara untuk menanggapi permintaan publik terkait kasus konten ilegal dan berbahaya secara online akan diidentifikasi.

Analisis pelaksanaan peraturan dan kebijakan negara terhadap masyarakat

Bagian ini berfokus pada bagaimana peraturan yang ada dapat mempengaruhi komunitas, dengan anggapan bahwa penegakan hukum pada konten daring yang ada saat ini dapat membahayakan kebebasan berekspresi dan secara tidak proporsional mempengaruhi komunitas rentan seperti LGBTQ+ dan kelompok minoritas agama.



Tinjauan Pustaka

Tinjauan pustaka akan menentukan praktik moderasi konten (*content moderation*) yang ada saat ini: untuk tujuan apa konten tersebut digunakan, siapa penggunanya, dan pada platform apa. Hal yang penting, bagian dari studi ini membantu mengidentifikasi isu-isu menonjol dalam moderasi konten secara keseluruhan, kontekstualisasi dalam isu penyebaran misinformasi, disinformasi, dan ujaran kebencian. Selain itu, praktik moderasi konten yang berbahaya dan/atau ilegal mungkin digunakan dengan mengorbankan kelompok marjinal, yang mengakibatkan berbagai bentuk penyensoran yang tidak disengaja dan disengaja yang dapat secara tidak proporsional berdampak pada komunitas yang paling tidak terlindungi. Tinjauan pustaka ini akan membahas hal-hal berikut: penilaian norma internasional dan regional pada konten ilegal dan berbahaya serta peran platform dalam memoderasi

Definisi Konten Ilegal dan Berbahaya berdasarkan Instrumen Hukum Internasional dan Regional

→ Instrumen Hukum Global

Salah satu tantangan dalam meregulasi konten daring secara efektif adalah menyelaraskan definisi dan/atau klasifikasi apa yang dimaksud dengan konten berbahaya atau ilegal. Sayangnya, tidak ada perjanjian dalam rezim hukum internasional yang secara spesifik mengatur konten ilegal dan berbahaya di internet. Hukum Hak Asasi Manusia Internasional (*The International Human Rights Laws, IHRL*) hanya secara implisit meregulasi konten ilegal melalui konsep yang luas seperti diskriminasi rasial. Maka, perbedaan kedua istilah tersebut hingga saat ini belum dieksplorasi dalam perjanjian internasional yang ada. Meski begitu, **Report of the Special Rapporteur (A/66/290) mengenai pengajuan dan perlindungan atas hak kebebasan berpendapat dan berekspresi** yang diserahkan kepada Majelis Umum pada sesi ke-66 tentang kebebasan berekspresi di internet, telah

menyajikan diskusi mengenai perbedaan konten ilegal dan berbahaya. Pelapor Khusus (*Special Rapporteur*) menggarisbawahi pernyataan tersebut sebagai berikut:

"... Ada beberapa perbedaan antara konten ilegal, yang harus dilarang oleh negara (AS) di bawah hukum internasional, seperti pornografi anak, dan konten yang dianggap berbahaya, menyinggung, tidak menyenangkan atau tidak diinginkan, tetapi negara tidak diwajibkan melarang maupun mengkriminalisasi. Dalam kasus ini, Pelapor Khusus percaya bahwa penting untuk membuat perbedaan yang jelas antara tiga bentuk ekspresi atau ungkapan: (a) ekspresi yang merupakan bentuk serangan di bawah hukum internasional dan dapat dituntut sebagai tindakan kriminal (b) ekspresi yang tidak dapat dihukum secara pidana namun dapat diberlakukan pembatasan dan gugatan perdata; dan (c) ekspresi yang tidak menimbulkan sanksi pidana maupun perdata, tetapi masih dapat menimbulkan kekhawatiran dalam hal toleransi, sopan santun, dan penghormatan terhadap orang lain. Kategori konten yang berbeda ini menimbulkan masalah prinsip dan memerlukan respon hukum serta teknologi secara berbeda."

Pelapor Khusus kemudian memberikan empat jenis ekspresi, yang termasuk dalam kategori utama, yaitu ekspresi yang termasuk dalam pelanggaran menurut hukum internasional serta dapat dituntut secara pidana. Keempat bentuk ekspresi tersebut termasuk (1) pornografi anak; (2) hasutan langsung dan publik untuk melakukan genosida; (3) pembelaan terhadap kebencian nasional, ras, atau agama yang merupakan hasutan untuk melakukan diskriminasi, permusuhan, atau kekerasan; dan (4)

hasutan untuk melakukan tindak terorisme. Sebagai catatan, mereka juga harus mematuhi pengujian tiga-bagian rekomendasi hukum yang jelas, mengejar tujuan yang sah, serta menghormati prinsip kebutuhan dan proporsionalitas.

Lebih lanjut, Pelapor Khusus menyebutkan bahwa bentuk ekspresi yang tidak termasuk dalam kategori pertama tidak bisa dipidanakan, termasuk hukum pencemaran nama baik yang bertujuan untuk melindungi reputasi seorang individu. Sebagaimana ditetapkan dalam Human Rights Council Resolution 16/12, Pelapor Khusus menekankan bahwa bentuk ekspresi berikut tidak boleh tunduk pada pembatasan, termasuk: diskusi kebijakan pemerintah dan debat politik; pelaporan pada hak asasi manusia, aktivitas pemerintahan, dan korupsi dalam pemerintahan; kampanye pemilu, demonstrasi perdamaian atau aktivitas politik, termasuk untuk perdamaian atau demokrasi; serta ekspresi opini dan perbedaan pendapat, agama atau kepercayaan, termasuk oleh orang-orang yang tergabung dalam kelompok minoritas atau rentan.

Selanjutnya, beberapa bentuk konten berbahaya secara umum masih kekurangan kejelasan dalam definisinya. Hal ini dikhawatirkan dapat membahayakan hak masyarakat atas kebebasan berpendapat dan berekspresi. Untuk mengantisipasi hal ini, beberapa instrumen dari PBB telah menyediakan definisi dari beragam konten berbahaya, seperti ujaran kebencian, disinformasi, dan misinformasi – yang mana menjadi fokus dalam penelitian ini.

The United Nation Strategy and Plan of Action on Hate Speech mendefinisikan ujaran kebencian sebagai ujaran yang menyerang atau menggunakan bahasa peyoratif maupun diskriminatif yang mengacu kepada perseorangan atau kelompok berdasarkan siapa mereka, dalam kata lain, berdasarkan agama, etnis, kewarganegaraan, ras, warna kulit, keturunan, gender, atau faktor identitas lain yang dimiliki. Meskipun telah didefinisikan sedemikian rupa, pengaturan mengenai ujaran kebencian juga harus memenuhi kriteria yang tercantum dalam ICCPR. Kriteria ini

mencakup beberapa poin yang akan dibahas lebih lanjut pada bagian selanjutnya.

Mengenai disinformasi dan misinformasi, Pelapor Khusus (A/HRC/47/25) mengklarifikasi definisi kedua istilah tersebut. Disinformasi didefinisikan sebagai informasi palsu yang disebarkan dengan sengaja guna menyebabkan kerugian sosial yang serius, sedangkan misinformasi didefinisikan sebagai penyebaran informasi palsu tanpa disengaja. Lebih lanjut, instrumen ini menekankan bahwa kedua istilah tersebut tidak dapat dipertukarkan.

Ketika terjadi kasus pencemaran nama baik, perangkat hukum yang ada saat ini tidak mengklarifikasikan definisi istilah tersebut. Namun, undang-undang pencemaran nama baik biasanya dibenarkan untuk melindungi reputasi perseorangan. Beberapa instrumen Pelapor Khusus kini telah meminta negara untuk mencabut semua undang-undang pidana pencemaran nama baik dan menggantinya dengan undang-undang perdata pencemaran nama baik. (Lihat, misalnya: Pelapor Khusus (E/CN.4/2000/63); Pelapor Khusus (E/CN.4/2001/64); Pasal 19, 2004). Pelapor Khusus (A/67/357) lebih lanjut menyatakan bahwa pencemaran nama baik tidak boleh ditetapkan pada kasus kritikan melawan pejabat publik. Seruan tersebut ditujukan untuk menghindari pembatasan berlebihan terhadap hak kebebasan berpendapat dan ekspresi masyarakat sebagaimana undang-undang pencemaran nama baik sering rentan disalahgunakan.

→ Instrumen Hukum Regional

● Instrumen Hukum Uni Eropa (European Union, EU)

Istilah 'konten ilegal dan berbahaya' pertama kali muncul secara resmi dalam sebuah dokumen yang diproduksi UE pada Oktober 1996, yaitu **The Communication on Illegal and Harmful Content on the Internet** (Price, 2002). Melalui dokumen tersebut, UE telah menunjukkan kepedulian mereka terhadap pentingnya membedakan konten ilegal

dan berbahaya. Mereka berargumen bahwa perbedaan kategori konten dapat “menimbulkan masalah prinsip yang sangat berbeda dan membutuhkan tanggapan hukum serta teknologi yang berbeda” (Komisi Eropa, 1996). Komisi Eropa (The European Commission) tidak merinci definisi dari konten ilegal. Meski begitu, beberapa contoh konten ilegal yang disebutkan melingkupi pelanggaran hak cipta, pencemaran nama baik, pelanggaran privasi, pornografi anak, penyebaran mater rasis, atau hasutan untuk kebencian rasial.

Lebih lanjut, Komisi Eropa juga menyadari bahwa definisi yang tepat untuk pelanggaran konten ilegal berbeda dari satu negara dengan negara lainnya (Komisi Eropa, 1996). Di sisi lain, konten berbahaya didefinisikan sebagai beragam bentuk materi yang dapat “menyinggung nilai dan perasaan orang lain”, termasuk konten menyuarakan opini politik, keyakinan agama, atau pandangan pada masalah ras. Namun demikian, dokumen tersebut mengelaborasi bahwa apa yang dianggap dapat berbahaya tergantung pada kekhasan budaya.

Di tahun yang sama, Komisi Eropa juga menerbitkan **Green Paper on the Protection of Minor and Human Dignity in Audio-visual and Information Services**. Meskipun dokumen ini tidak secara spesifik membedakan antara ‘ilegal dan berbahaya’, dokumen ini menawarkan varian ‘materi yang dilarang untuk semua oleh negara anggota tertentu’ dan ‘materi tertentu yang dapat berdampak pada perkembangan fisik serta mental anak di bawah umur’. Utamanya, kategori pertama terdiri dari pornografi anak, kekerasan serampangan yang ekstrim, dan hasutan untuk kebencian rasial atau lainnya, diskriminasi, dan kekerasan. Di sisi lain, kategori kedua termasuk reklame atau iklan.

The European Commission’s Action Plan for the European Union for a Safer Use of the Internet pada tahun 2007 menyatakan bahwa konten ilegal terkait dengan beragam isu yang luas tersebut

selayaknya merakit bom, yang dapat mengancam keamanan nasional, pornografi anak, hasutan terhadap kebencian rasial, dan pencemaran nama baik. Sebagai perbandingan, konten berbahaya adalah konten yang diizinkan namun juga memiliki batas edar (misalnya, hanya untuk orang dewasa) dan konten yang berpotensi menyinggung beberapa pengguna, meskipun publikasinya tidak dibatasi untuk menjamin kebebasan berbicara.

Pada pertengahan Desember 2020, Komisi Eropa mengungkapkan usulan Digital Services Act (DSA). Usulan DSA ini berusaha menyiapkan regulasi yang efisien terkait layanan digital inovatif yang efisien untuk pasar internal, mempromosikan keamanan daring dan perlindungan hak-hak dasar, serta membangun tata kelola efektif dalam pengawasan penyedia layanan perantara (Komisi Eropa, 2020). Disahkan oleh UE pada April 2022, DSA dapat digunakan sebagai referensi dalam membahas perbedaan antara konten ilegal dan berbahaya.

Dalam nota penjelasan, mengindikasikan bahwa konsultasi multi-stakeholder untuk proposal menghasilkan kesepakatan umum di antara para pemangku kepentingan, yaitu bahwa: "konten 'berbahaya' tidak boleh didefinisikan dalam Undang-undang Layanan Digital dan tidak boleh tunduk pada kewajiban penghapusan, sebagaimana ini adalah area sensitif dengan implikasi yang berat pada perlindungan kebebasan berekspresi." Dengan demikian, usulan DSA hanya mendefinisikan 'konten ilegal' dan tidak memuat definisi 'konten berbahaya'. Pasal 2 huruf (g) pada usulan DSA menyatakan definisi 'konten ilegal' sebagai berikut:

"... 'konten ilegal' berarti informasi apapun, yang, secara independen maupun merujuk pada suatu aktivitas, meliputi penjualan produk atau penyediaan layanan yang tidak sesuai dengan hukum Union (Union Law) atau hukum negara anggota (the Law of a Member State), terlepas dari pokok permasalahan atau sifat hukum tersebut."

Tidak adanya garis yang jelas antara perbedaan konten ilegal dan konten berbahaya dalam usulan DSA dinilai membuat definisi konten berbahaya terlalu luas dan lebih meluaskan keleluasaan penyedia layanan perantara dalam mengartikannya (Branden et al, 2021).

Satu studi yang dilakukan oleh Parlemen Eropa berjudul 'Reform of the EU Liability Regime for Online Intermediaries' juga telah memberikan dasar yang cukup kuat untuk membedakan konten ilegal dan konten berbahaya. Konten ilegal didefinisikan sebagai berbagai macam informasi yang tidak sesuai atau tunduk pada undang-undang UE dan nasional, seperti ujaran kebencian, hasutan melakukan kekerasan, materi kekerasan pada anak, dan pornografi balas dendam. Di sisi lain, konten berbahaya adalah informasi yang tidak sepenuhnya termasuk dalam larangan hukum tetapi mungkin memiliki efek berbahaya, seperti perundungan siber (*cyberbullying*) dan mis/disinformasi (Madiega, 2020).

● **Instrumen Hukum Asosiasi Negara Asia Tenggara (ASEAN)**

Tidak ada instrumen yang mengikat secara tunggal yang mengatur konten ilegal dan berbahaya di ASEAN. Maka dari itu, perbedaan antara konten ilegal dan berbahaya juga tidak bisa ditemukan. Selama ini, ASEAN baru menghasilkan beberapa jenis *soft law*, seperti deklarasi, untuk mengatur bentuk-bentuk perbuatan yang dapat diklasifikasikan sebagai konten berbahaya.

Di tahun 2017, ASEAN mengadopsi **The ASEAN Declaration to Prevent and Combat Cybercrime**. Deklarasi ini memuat komitmen negara anggota ASEAN untuk berkolaborasi dalam upaya untuk mencegah dan memberantas kejahatan siber. Lebih lanjut, deklarasi ini juga mengakui pentingnya harmonisasi hukum yang terkait dengan kejahatan siber dan barang bukti elektronik pada level regional.

Selanjutnya, Konferensi Menteri ASEAN yang Bertanggungjawab untuk Informasi ke-14 (*the 14th Conference of the ASEAN Ministers Responsible for Information, AMRI*) pada tahun 2018 mengeluarkan *Framework and Joint Declaration to Minimize the Harmful Effects of Fake News*. Dokumen tersebut mengandung beberapa strategi yang dapat digunakan oleh negara untuk memerangi berita palsu, mulai dari penyusunan undang-undang, norma, dan/atau pedoman nasional dan peningkatan keterlibatan masyarakat sipil. Namun, definisi dari berita palsu juga belum disebutkan.

Sebagai kesimpulan, beberapa instrumen hukum yang dijelaskan di atas mengilustrasikan pentingnya membedakan konten 'ilegal' dan konten 'berbahaya' dalam mengatur konten di media sosial, dengan pertimbangan kemungkinan perbedaan dalam menangani dan mengelola konten ilegal dan berbahaya secara praktik antara negara dan platform (Madiega, 2020). Sebagai contoh, pemangku kepentingan menangani 'konten ilegal' dengan cara menghapusnya dari platform media sosial. Kemudian, kemungkinan tersangka dapat diproses berdasarkan hukum yang ada. Sementara, 'konten berbahaya' dapat diproses berdasarkan pedoman komunitas pada masing-masing platform media sosial (Vogelezang, 2020). Dalam konteks regional, badan independen juga dapat dibentuk untuk mengatasi ketidakharmonisan dari pedoman komunitas antara platform dan regulasi antar negara (De Streel, 2020). Lebih lanjut, perbedaan antara konten ilegal dan konten berbahaya juga dapat bermanfaat dalam menyeimbangan hak-hak pengguna internet sembari menjaga platform daring bertanggung jawab kepada regulator (Vogelezang, 2020).

Norma dan Standar Internasional dalam Meregulasi Konten Ilegal dan Berbahaya

Tidak dapat dipungkiri bahwa kebebasan berekspresi adalah bagian dari hak asasi manusia, sebagaimana tercermin dalam Pasal 19 dari the Universal Declaration of Human Rights (UDHR). Pasal 19 dalam UDHR tersebut berbunyi sebagai berikut:

“Setiap orang berhak atas kebebasan untuk memiliki dan mengeluarkan pendapat; dalam hal ini termasuk kebebasan menganut pendapat tanpa mendapat gangguan, dan untuk mencari, menerima dan menyampaikan keterangan-keterangan dan pendapat dengan cara apa pun dan dengan tidak memandang batas-batas.”

UDHR tidak secara eksplisit merujuk pada kebebasan berekspresi di internet. Hal tersebut adalah lumrah karena sulit dibayangkan bahwa para perancang pada masa itu akan memahami situasi dan kondisi saat ini. Namun, Pengadilan Hak Asasi Manusia Eropa (the European Court of Human Rights) merujuk pada interpretasi pada Pasal 10 UDHR, yang berbunyi ‘dilihat dari kondisi saat ini’. Ranah internet tidak diragukan lagi telah menjadi bagian penting dalam perihal berkomunikasi saat ini (Benedek & Kettemann, 2013). Dengan demikian, kebebasan berekspresi di internet juga bisa termasuk dalam lingkup pasal 19 UDHR.

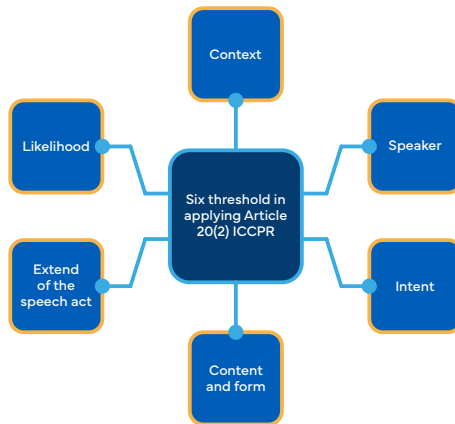
Internet dan sosial media diciptakan sebagai media komunikasi bagi manusia satu sama lain. Dalam berkomunikasi, orang tertentu mengekspresikan berbagai ekspresi dalam dirinya. Tidak seperti di dunia nyata, ekspresi orang dalam berkomunikasi di internet muncul dalam bentuk ‘konten’. Pengungkapan ekspresi di ranah internet juga perlu diatur. Oleh karena itu, diperlukan adanya regulasi tentang konten ilegal dan konten berbahaya di internet. Tantangan terbesar dari penyusunan regulasi

ini adalah bagaimana merumuskan peraturan untuk melindungi masyarakat dari konten berbahaya sembari tetap menjaga kebebasan berekspresi. Maka dari itu, perlu ada pembedaan antara konten ilegal dan konten berbahaya (Burns, 2020).

Hukum Hak Asasi Manusia Internasional (the International Human Rights Law, IHRL), membahas masalah konten ilegal dan konten berbahaya. Dalam hal ini, IHRL memiliki dua fungsi. Fungsi pertama adalah untuk mengamankan pelarangan bentuk ekspresi tertentu, sedangkan bentuk kedua adalah menjadi pembatas jenis konten yang mungkin dilarang oleh negara (Sander, 2021). Sebagaimana disebutkan di atas, The Universal Declaration on Human Rights Law (UDHR) adalah dokumen kunci yang mengilhami banyak mekanisme hukum yang menangani hak asasi manusia, termasuk perjanjian internasional yang paling relevan yang mengatur hak asasi manusia. Deklarasi tersebut menjadi acuan bagi banyak negara dalam merumuskan peraturan nasional hak asasi manusianya, termasuk peraturan yang berkaitan dengan konten ilegal dan berbahaya. UDHR juga dijabarkan lebih lanjut dalam instrumen hukum internasional lainnya seperti the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD), the International Covenant on Economic, Social and Cultural Rights, (ICESCR), dan the International Covenant on Civil and Political Right (ICCPR) yang juga telah diratifikasi Indonesia melalui Undang-Undang Nomor 12 Tahun 2005 tentang pengesahan ICCPR.

Beberapa konvensi internasional di atas setidaknya dapat menjadi dasar untuk mengatur konten ilegal dan berbahaya/ Pasal 4 ICERD mewajibkan negara-negara untuk mengkriminalisasi semua penyebaran gagasan berdasarkan superioritas atau kebencian rasial, hasutan untuk diskriminasi rasial, serta semua tindakan kekerasan atau hasutan untuk melakukan tindakan tersebut terhadap kelompok ras atau etnis mana pun. Ketentuan serupa dapat ditemukan dalam Pasal 20 Ayat 2 ICCPR yang menetapkan bahwa setiap negara diwajibkan melarang 'setiap advokasi kebencian nasional, ras, atau agama yang merupakan hasutan untuk

melakukan diskriminasi, permusuhan, atau kekerasan'. Berdasarkan ketentuan masing-masing, Pelapor Khusus Perserikatan Bangsa-Bangsa (PBB) tentang kebebasan berekspresi (A/74/486) memperluas cakupan ujian kebencian di luar 'kebencian nasional, rasial, atau agama', meluas ke tindakan merugikan atas dasar 'ras, warna kulit, jenis kelamin, bahasa, agama, pendapat politik atau lainnya, asal kebangsaan atau sosial, properti, kelahiran atau status lainnya, termasuk asal atau identitas orang asli, kecacatan, status migran atau pengungsi, orientasi seksual, identitas gender atau status interseks'. The Rabat Plan of Action, yang diadopsi oleh sekelompok pakar hak asasi manusia tingkat tinggi, menyarankan agar negara-negara bagian melakukan tes ambang batas enam bagian ketika menerapkan Pasal 20(2) ICCPR. Keenam ambang batas tersebut meliputi konteks, pembicara, niat, isi dan bentuk, jangkauan tindak tutur, dan kemungkinan. Di bawah ini merupakan diagram serta penjelasan singkat masing-masing ambang batas yang disebutkan di atas.



Gambar 1. Tes Ambang Enam Bagian Penerapan Pasal 20(2) ICCPR
(Sumber: The Rabat Plan of Action 2011)

● **Konteks (Context):** Pidato / ujaran yang bersangkutan harus dianalisis dalam konteks sosial dan politik yang lazim pada saat tindakan itu dilakukan dan disebarluaskan.

- **Pembicara (*Speaker*):** Pelaku yang membuat atau menyebarkan pidato / ujaran harus dianalisis terkait posisi atau status mereka dalam masyarakat.
- **Niat (*Intent*):** Kelalaian dan kecerobohan tidak cukup untuk memenuhi ambang batas. Maka, hubungan segitiga antara obyek dan subyek tindak tutur serta pemirsanya harus dianalisis lebih lanjut.
- **Isi dan Bentuk (*Content and form*):** Isi dan bentuk pidato / ujaran harus dianalisis untuk dapat memeriksa maksud dan risiko bahaya dari pidato / ujaran tersebut.
- **Jangkauan Tindak Tutur (*Extent of the speech act*):** Jangkauan tindak tutur harus dianalisis, termasuk besaran tuturan dan ukuran khalayaknya.
- **Kemungkinan (*Likelihood*):** Beberapa tingkat risiko bahaya yang timbul dari ucapan harus diidentifikasi.

Selanjutnya, Pasal 19 ICCPR menjadi dasar bagi hak dan kewajiban berekspresi dan tanggung jawab negara dalam melindungi kebebasan tersebut. The UN Human Rights Committee (HCR) juga menjelaskan bahwa negara-negara tersebut harus mematuhi standar yang sudah dijabarkan pada Pasal 19(3) ICCPR, yang berbunyi seperti berikut:

'Pelaksanaan [hak kebebasan berekspresi] disertai dengan tugas dan tanggung jawab khusus. Oleh karena itu dapat dikenakan pembatasan tertentu, tetapi hal ini hanya akan seperti yang ditentukan oleh hukum serta diperlukan: (a) untuk menghormati hak atau reputasi orang lain; (b) untuk melindungi keamanan nasional atau ketertiban umum (ordre public), atau kesehatan atau moral masyarakat.'

Berdasarkan General Comment No. 34 on Article 19 of the ICCPR, pengaturan terkait kebebasan berekspresi dalam Pasal 19 ICCPR juga mencakup semua bentuk audio-visual serta mode berekspresi berbasis elektronik dan internet. Namun, The UN Human Rights Committee Joint Declaration on Freedom of Expression lebih lanjut mencatat bahwa pendekatan regulasi telekomunikasi saat ini tidak mudah diterapkan dalam konteks internet (Pasal 19, 2021).

Laporan dari Pelapor Khusus mengenai kemajuan dan perlindungan hak kebebasan berpendapat dan berekspresi (A/HRC/38/35) menyatakan bahwa negara tidak boleh membatasi hak memiliki pendapat tanpa campur tangan. Dalam menerapkan Pasal 19(3) ICCPR, pembatasan oleh negara atas kebebasan berekspresi harus dapat memenuhi ketentuan berikut:

- **Keabsahan (*Legality*):** pembatasan tersebut harus diadopsi melalui proses hukum biasa dan membatasi kebijaksanaan pemerintah dengan membedakan antara ekspresi yang sah dan yang melanggar hukum dengan “ketepatan dengan kadar sesuai”.
- **Kebutuhan dan Kesepadanan (*Necessity and Proportionality*):** negara harus menunjukkan pembatasan tersebut memberikan beban paling ringan dalam rangka menjamin hak asasi, serta benar-benar melindungi, atau memberikan perlindungan kepada kepentingan sah negara yang memerlukan perhatian.
- **Legitimasi (*Legitimacy*):** pembatasan apapun harus melindungi kepentingan yang ditentukan dalam Pasal 19(3), termasuk hak atas reputasi orang lain, keamanan nasional atau ketertiban umum, atau kesehatan atau moralitas.

Selain Pasal 19(3) ICCPR, kondisi kumulatif di atas harus dipenuhi oleh negara dalam menerapkan Pasal 20(2) ICCPR yang mewajibkan negara melarang “penganjuran kebencian nasional, ras, atau agama yang merupakan hasutan melakukan diskriminasi, permusuhan, atau kekerasan”.

Laporan Pelapor Khusus mengenai pemajuan dan perlindungan hak kebebasan berpendapat dan berekspresi (A/74/486) lebih lanjut menjelaskan beberapa jenis peraturan yang sering diadopsi negara tapi tidak memenuhi persyaratan yang tercantum dalam Pasal 19(3) ICCPR. Salah satunya adalah undang-undang anti penodaan agama.

Selain instrumen hukum yang mengikat di atas, PBB juga mengeluarkan the UN Strategy and Plan of Action on Hate Speech. Meskipun sifatnya dapat digolongkan sebagai soft law, dokumen tersebut memberikan definisi ujaran kebencian yang dapat digunakan sebagai dasar diskusi mengenai konten ilegal dan berbahaya. The UN Strategy and Plan of Action on Hate Speech mendefinisikan ujaran kebencian sebagai segala bentuk komunikasi dalam ucapan, tulisan, atau pun perilaku, yang menyerang atau menggunakan bahasa yang merendahkan atau diskriminatif dengan mengacu pada seseorang atau kelompok berdasarkan siapa mereka, dengan kata lain, berdasarkan agama, suku, kebangsaan, ras, warna kulit, keturunan, gender atau faktor identitas lainnya. Dokumen tersebut juga menegaskan bahwa ujaran kebencian termasuk sebagai tindakan yang tidak dilarang tetapi dapat merugikan atau membahayakan.

Baru-baru ini, Laporan Pelapor Khusus tentang pemajuan dan perlindungan hak kebebasan berpendapat dan berekspresi (A/HRC/47/25) memberikan definisi disinformasi yang dapat digunakan sebagai dasar tambahan untuk diskusi mengenai konten ilegal dan berbahaya. Disebutkan bahwa disinformasi dipahami sebagai informasi palsu yang disebarkan dengan sengaja guna menimbulkan kerugian sosial yang serius, dan misinformasi sebagai penyebaran informasi palsu tanpa disadari. Lebih lanjut, dijelaskan juga bahwa beberapa bentuk disinformasi dapat menjadi hasutan kebencian, diskriminasi, dan kekerasan, yang dilarang menurut hukum internasional. Selain itu, dokumen tersebut menegaskan kembali bahwa penting untuk mengklarifikasi konsep disinformasi dan misinformasi dalam kerangka IHRL.

Laporan Pelapor Khusus tentang pemajuan dan perlindungan hak kebebasan berpendapat dan berekspresi tersebut (A/HRC/38/35) juga menghasilkan beberapa rekomendasi bagi negara-negara dalam melindungi hak berpendapat dan berekspresi di internet, yaitu bahwa: (1) negara harus mencabut undang-undang yang mengkriminalisasi atau membatasi ekspresi secara berlebihan; (2) negara hanya harus berusaha membatasi konten sesuai dengan perintah dari otoritas peradilan yang independen dan tidak memihak serta sesuai dengan proses hukum dan kondisi kumulatif keabsahan, kebutuhan, dan legitimasi; (3) negara tidak boleh mengadopsi model regulasi di mana pemerintah selain otoritas yudisial, menjadi penengah ekspresi yang sah. Mengenai perusahaan Teknologi Informasi dan Komunikasi (TIK), laporan tersebut memberi rekomendasi bahwa mereka harus mengakui standar global yang berwenang guna memastikan kebebasan berekspresi pada platform mereka adalah hukum hak asasi manusia, dan bukan berbagai hukum negara bagian untuk kepentingan perusahaan.

Selanjutnya, Budapest Convention on Cybercrime merupakan perjanjian internasional pertama yang secara khusus membahas masalah internet dan kejahatan komputer. Meskipun tidak membedakan konten ilegal dan konten berbahaya, konvensi ini mengatur beberapa tindakan di mana para pihak terkait harus mengambil tindakan legislatif dan tindakan lain yang mungkin dibutuhkan untuk ditetapkan sebagai tindak pidana menurut hukum domestik mereka. Tindakan yang dimaksud mencakup pelanggaran yang berkaitan dengan pornografi anak dan pelanggaran terkait pelanggaran hak cipta dan hak-hak lainnya yang terkait.

Berdasarkan beberapa instrumen hukum yang disebutkan, dapat disimpulkan bahwa tidak ada perjanjian internasional yang secara khusus mengatur konten ilegal dan berbahaya di internet. Namun, instrumen hukum internasional yang dapat menjadi dasar pengaturan konten berbahaya sebenarnya sudah ada. Sayangnya, pengaturan tersebut masih tersebar di beberapa perjanjian internasional, bahkan dalam beberapa dokumen '*soft law*'. Setidaknya diperlukan instrumen hukum internasional

yang komprehensif yang mengatur konten berbahaya untuk merespons secara efektif peningkatan penggunaan media sosial.

Sebagai tambahan, meski tidak harus terdaftar dalam naungan instrumen hukum internasional, memastikan transparansi dalam mengatur konten dianggap lebih diperlukan. Beberapa penelitian berpendapat bahwa peningkatan transparansi moderasi konten sangat penting untuk memastikan pemantauan konten dan praktik moderasi yang lebih efektif (WEF, 2021; Gorwa et al., 2020; APC, 2018; & De Gregorio, 2020). Selain itu, Pelapor Khusus PBB menekankan pentingnya penegakan kerangka kerja hak asasi manusia internasional dalam memoderasi konten online (United Nations Human Rights Office of the High Commissioners [OHCHR], 2018). Pada aspek ini, model moderasi konten Uni Eropa (UE), melalui EU Digital Service Act (DSA), juga menekankan hak-hak dasar dan prinsip proporsionalitas dalam pengaturan konten di dalam yuridiksinya (European Parliament, 2020).

Sayangnya, terdapat kesenjangan yang jelas antara praktik terbaik dengan tindakan yang diambil negara saat mereka menggunakan pendekatan yang lebih opresif. Pendekatan ini sering mengabaikan prinsip hak asasi manusia. Sebagai akibatnya, hal ini dapat digunakan untuk membatasi kebebasan berpendapat serta berpotensi digunakan sebagai alat penyensoran (Banchik, 2020). Misalnya, NetzDg Jerman memaksa penghapusan konten berdasarkan ketentuan hukum pidananya, yang mana hal ini meningkatkan kekhawatiran tentang kebebasan berekspresi dengan konsep luas ketentuan 'penistaan agama' dan 'ujaran kebencian' (Pasal 19, 2017). Yang lebih memprihatinkan, beberapa negara bagian lainnya seperti India, Kenya, Malaysia, Filipina, Rusia, Turki, dan Venezuela, memberlakukan peraturan konten serupa (WEF, 2021). Hal ini menunjukkan bagaimana negara secara aktif memilih untuk menerapkan pendekatan mekanisme hukuman – suatu pendekatan yang berbeda – bahkan bertentangan dengan norma-norma yang diusulkan pada peraturan konten. Oleh karena itu menyoroti kesenjangan yang mencolok dan mendasar antara norma-norma yang diusulkan pada peraturan konten,

seperti yang ditemukan dalam beberapa studi dan saran dari pelapor, dan praktik aktual yang diterapkan oleh negara, seperti yang ditunjukkan dalam adopsi regulasi konten hukuman secara luas.

Platform and Moderasi Konten

Platform media sosial (seperti Facebook, Instagram, Twitter, dan YouTube) memiliki, serta mendapatkan keuntungan finansial dari sejumlah besar pengguna dan data yang mereka hasilkan (Maulana, 2021). Selain itu, mereka juga merupakan pemimpin utama dalam praktik pengguna dan komunitas di platform mereka. Dikarenakan platform media sosial memiliki kekuasaan yang sangat besar, publik harus tetap meminta pertanggungjawaban platform dalam melindungi mereka (Doctorow, 2021b), terutama jika tujuan yang dinyatakan oleh para perusahaan ini diharapkan dapat dipercaya oleh publik.

Meski demikian, perkembangan terkini dan literatur yang ada menyarankan perusahaan-perusahaan ini agar berusaha keras guna memastikan profitabilitas dibanding kepentingan pengguna (misalnya, Aschoff, 2020; Marcetic, 2021; Oates, 2020). Gagasan mengenai pertumbuhan, keterlibatan, dan skala adalah penting untuk platform media sosial guna menarik pengiklan dan para investor. “Pengguna Aktif Harian” dan “Pengguna Aktif Bulanan” adalah indeks penting untuk kesuksesan perusahaan media sosial, dan seringnya di atas nilai-nilai lain seperti pengembangan sosial, keamanan, kualitas, dan interaksi yang bermakna (Zulli et al., 2020, hlm. 1195). Ekstraksi data, modifikasi perilaku, serta eksternalisasi biaya menjadi tidak mengejutkan karena perusahaan teknologi, sebagaimana perusahaan lainnya, didorong oleh kebutuhan profitabilitas jangka panjang (Morozov, 2019).

Pendekatan yang berorientasi pada keuntungan (*profit-oriented*) dapat berdampak negatif pada praktik moderasi konten karena platform akan terus mencoba meningkatkan keterlibatan pengguna dan waktu layar (screen time) pada situs mereka. Reviglio dan Astoti (2020, hlm. 4) menemukan bahwa polarisasi sosial dan politik yang sering dijumpai di

media sosial sebenarnya merupakan bagian dari model bisnis mereka. Mereka melakukannya dengan membuat umpan konten yang terlalu homofili sehingga menghasilkan kelompok yang terpolarisasi dan konten yang memecah belah. Dikarenakan kelompok ini bermuatan emosi dan politik, mereka cenderung mempertahankan keyakinan mereka dan menyerang orang lain. Salah satu contoh yang bisa menjelaskan hal ini adalah kasus Pemilihan Gubernur (Pilgub) Jakarta 2017, ketika pendukung dua kandidat, Wakil Gubernur aktif berubah menjadi Gubernur Basuki Tjahaja Purnama (Ahok) dan Anies Baswedan, berpartisipasi dalam apa yang disebut Merylina Lim (2017) sebagai nasionalisme suku. Didorong oleh penanda ras dan agama, para pendukung kedua kandidat mengekstrapolasi perpecahan mereka dari aktivitas politik luring ke dunia daring dan menciptakan kantong-kantong algoritmik.

Lim (2017) mengemukakan algoritma kantong (enclaves) merupakan komunitas imajiner berbasis tekno-sosial yang dibangun ketika kelompok individu, difasilitasi interaksi konstan mereka dengan algoritma, mencoba membentuk rasa identitas (yang dirasakan) secara daring, di mana mereka akan mempertahankan keyakinan mereka dan menyerang orang lain. Meski algoritma itu sendiri tidak menentukan pembuatan kantong ini, tetapi kantong tersebut terbentuk dengan sendirinya berdasarkan interaksi algoritma-pengguna. Lebih jauh, Lim (2017, hlm. 13) berpendapat bahwa kantong-kantong ini dapat terlihat dengan sangat jelas dalam lingkup sosial yang sudah terfragmentasi antara "Cina-Kristen, hijabi (baik pro dan anti-Ahok), dan pribumi".

Kritik lain terhadap mekanisme pengaturan mandiri dari platform adalah kurangnya transparansi dalam metode pengaturan konten. Sejumlah besar konten yang dihasilkan setiap menit mendorong platform untuk menggunakan alat pemantau canggih, seperti algoritma dan mesin AI, untuk memantau dan mengatur konten mereka. Algoritma sering digunakan untuk menandai konten yang melanggar standar komunitas. Jika dikembalikan lagi ke perangkat masing-masing, algoritma dapat menandai dan menghapus terlalu banyak konten atau malah membiarkan

ujaran negatif tetap utuh karena mereka tidak cukup canggih dalam menguraikan bahasa minoritas dan dialek daerah (Sinpeng & Martin, 2021).

Gorwa dkk. (2020) juga mengamati bahwa moderasi otomatis menggunakan alat pemantauan canggih milik platform sering mengakibatkan kesulitan dalam pengauditan proses, dan memiliki tantangan keadilan yang rumit (karena beberapa pandangan dan nilai dari kelompok tertentu diistimewakan) mendepolitisasi proses politik dalam moderasi konten platform. Mereka berpendapat bahwa yang penting bukan bagaimana meningkatkan efisiensi platform dalam melacak dan menghapus konten berbahaya, tetapi proses bagaimana mereka menentukan suatu konten sebagai hal yang berbahaya.

Selain itu, Association for Progressive Communications (APC, 2018) menemukan bahwa banyak platform memiliki perjanjian yang tidak diungkapkan dan tidak transparan (undisclosed and non-transparent agreements) dengan negara dalam memoderasi bahkan menghapus konten dari platform. Perjanjian tidak transparan ini disebut guna menghindari lembaga demokrasi dan memfasilitasi penyensoran terhadap pidato atau ujaran yang sah. Salah satu contoh yang menonjol adalah kerja sama Israel dan Facebook dalam menangani konten 'hasutan' di mana istilah tersebut didefinisikan secara samar, sedangkan salah satu interpretasinya adalah konten yang menentang dan mengkritik kebijakan Israel (APC, 2018).

Ada juga masalah ketenagakerjaan dalam praktik moderasi konten platform. Selain menggunakan alat pemantauan canggih, platform juga memberlakukan praktik manual berbasis tenaga kerja untuk memantau konten dalam situs mereka. Facebook mempekerjakan 15.000 pekerja, Google mempekerjakan sekitar 10.000 pekerja untuk mengawasi berbagai produk mereka – seperti YouTube – dan Twitter memiliki sekitar 1.500 moderator. Sayangnya, sebagian besar moderator konten ini adalah pekerja outsourcing, yang mengakibatkan beberapa masalah seperti tenaga kerja yang dibayar rendah dan kondisi kerja yang mengerikan – bahkan traumatik (Barrett, 2020).



BAB I

Regulasi pada Konten Berbahaya Daring di Indonesia



Bab pertama ini akan memetakan dan menganalisis regulasi di Indonesia terkait konten ilegal dan konten berbahaya. Pertama, kami mengkaji dasar konstitusional terkait regulasi konten ilegal dan berbahaya, terutama yang terkait dengan hak dan batasan konstitusional. Kedua, penulis menganalisis klasifikasi konten ilegal dan berbahaya berdasarkan regulasi yang ada. Ketiga, solusi serta mekanisme yang berlaku terhadap penanganan konten ilegal dan berbahaya akan dijelaskan. Keempat, kami menyoroti tanggung jawab platform media sosial berdasarkan regulasi yang ada. Terakhir, kami merangkum implikasi hukum regulasi konten di Indonesia.

Dasar Konstitusional Pengaturan "Konten Ilegal dan Berbahaya" di Indonesia

Wacana mengenai konten ilegal dan konten berbahaya tidak terlepas dari pembahasan regulasi dan hak asasi manusia. Sebagai negara hukum, peraturan perundang-undangan Indonesia memegang peranan penting dalam mengatur berbagai aspek kehidupan masyarakat, termasuk cara mereka berinteraksi dengan orang lain. Oleh karena itu, peraturan negara menjadi dasar utama dalam mengidentifikasi konten ilegal dan berbahaya serta cara menanganinya.

Lebih lanjut, pembahasan mengenai konten yang memiliki potensi bahaya tidak terlepas dari konteks perlindungan hak asasi manusia. Pengaturan konten tersebut berpotensi membatasi hak asasi warga negara. Mengenai regulasi konten berbahaya, Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (UUD 1945) secara eksplisit telah

menjamin perlindungan hak asasi manusia, termasuk kebebasan berekspresi dan kebebasan informasi, sebagai berikut:

Pasal 28

Kebebasan berserikat dan berkumpul, mengeluarkan pikiran dengan lisan, dan sebagainya ditetapkan dalam Undang-Undang.

Pasal 28E

(3) *Setiap orang berhak atas kebebasan berserikat, berkumpul, dan mengeluarkan pendapat.*

Pasal 28F

Setiap orang berhak untuk berkomunikasi dan memperoleh informasi untuk mengembangkan pribadi dan lingkungan sosialnya, serta berhak untuk mencari, memperoleh, memiliki, menyimpan, mengolah, dan menyampaikan informasi dengan menggunakan segala jenis saluran yang tersedia.

Ketentuan tersebut, khususnya Pasal 28E dan Pasal 28F yang termasuk dalam amandemen kedua UUD 1945, diambil dari Pasal 19, Pasal 20, dan Pasal 21 Universal Declaration of Human Rights (UDHR), yang didalamnya terdapat persamaan kata. Hal itu ditegaskan para pembuat konstitusi dalam amandemen ke-2 UUD 1945 tahun 2000 (Mahkamah Konstitusi, 2010).

Untuk menjamin pelaksanaannya, Pasal 28I UUD 1945 juga memberikan kewajiban konstitusional kepada negara, khususnya pemerintah, untuk melindungi, memajukan, menegakkan, dan memenuhi hak asasi manusia. Secara umum dapat dipahami bahwa negara memiliki tanggung jawab utama untuk melindungi hak asasi manusia. Dengan demikian, negara harus memberikan perlindungan dan jaminan untuk mewujudkan hak asasi

warga negara (Nickel, 1993). Peran negara juga ditekankan, misalnya, dalam the International Covenant on Civil and Political Rights, yang memberikan tanggung jawab umum negara untuk melindungi hak asasi manusia. Perlindungan hak asasi manusia yang diberikan negara dapat dilakukan melalui proses politik dan perumusan perangkat hukum (Evans, 2009).

Selanjutnya, perangkat hukum yang dibentuk juga berfungsi sebagai sarana yang sah dalam membatasi hak asasi manusia. Di Indonesia, UUD 1945 mengatur bahwa penikmatan hak asasi manusia hanya dapat dibatasi melalui UU. Rinciannya tercantum dalam Pasal 28I UUD 1945 sebagai berikut:

*Dalam melaksanakan hak dan kebebasannya, setiap orang harus tunduk pada **pembatasan yang diberikan oleh Undang-undang** untuk menjamin pengakuan dan penghormatan atas hak dan kebebasan orang lain semaksimal mungkin dan untuk **memenuhi tuntutan yang adil sesuai dengan pertimbangan moral, agama, nilai, keamanan, dan ketertiban umum dalam masyarakat yang demokratis.***

Pandangan Konstitusi Indonesia mengenai pembatasan hak asasi manusia juga diilhami oleh hukum internasional, khususnya UDHR.⁵ Ketentuan tersebut dijabarkan lebih lanjut dalam the International Covenant on Civil and Political Rights (ICCPR).⁶ Maka, sangat penting untuk memastikan bahwa instrumen hukum yang dirumuskan negara dapat memberikan perlindungan hukum yang kuat dan jaminan hak asasi manusia, termasuk kebebasan berbicara dan berekspresi yang sering diserang pada era digital sekarang ini.

⁵ Perlu dicatat bahwa Universal Declaration of Human Rights (UDHR) mengilhami perumusan ketentuan hak asasi manusia dalam UUD 1945. Hak tersebut tertuang dalam Naskah Komprehensif Perubahan UUD 1945 Buku VIII. Lihat Pasal 29 Ayat 2 Deklarasi Universal Hak Asasi Manusia.

⁶ Lihat misalnya Pasal 11 Ayat 3, Pasal 18 Ayat 3, Pasal 19 Ayat 3 Huruf (b), dan Pasal 22 Ayat 3. Konvensi regional lainnya juga memiliki ketentuan serupa seperti the European Convention of Human Rights (Lihat Pasal 8 Ayat 2, Pasal 9 Ayat 2, dan Pasal 10 Ayat 2), the American Convention of Human Rights (Lihat Pasal 12 Ayat 3, Pasal 13 Ayat 2 Huruf (b), dan Pasal 15). Selanjutnya, Indonesia juga telah meratifikasi ICCPR melalui UU No.12 Tahun 2005.

Analisis terhadap Undang-undang tersebut juga akan mengelaborasi apakah definisi atau penjelasan mengenai konten yang melanggar hukum dan merugikan sudah cukup jelas dan apakah Undang-undang terkait telah memenuhi syarat konstitusional, termasuk kepedulian terhadap moralitas, nilai-nilai agama, keamanan, dan ketertiban umum. Selanjutnya, cara regulasi menangani konten ini juga akan dianalisis. Selain itu, kami akan mengeksplorasi berbagai keputusan pengadilan yang berhubungan dengan Undang-undang terkait untuk memahami lebih lanjut bagaimana pengadilan menafsirkan ketentuan konten ilegal dan berbahaya. Sangat penting untuk memahami seutuhnya bagaimana pengadilan menerapkan hukum dalam memutuskan kasus terkait konten-konten ini.

Klasifikasi Konten Ilegal dan Berbahaya di dalam Regulasi Indonesia

Di Indonesia, UU ITE dapat dikatakan sebagai basis legal primer untuk meregulasi konten online. Undang-undang ini pada dasarnya mengatur berbagai aspek dunia maya, termasuk konten digital. Undang-undang ini juga secara luas digunakan untuk menangani berbagai kejahatan yang berkaitan dengan kejahatan dunia maya, khususnya yang terasosiasi dengan konten ilegal dan berbahaya. Meskipun UU ITE tidak secara eksplisit mendefinisikan “konten”, hal ini dapat diinterpretasikan sebagai informasi elektronik dan dokumen elektronik menurut UU ITE. Keduanya dapat dideskripsikan sebagai berikut:⁷

⁷ Pasal 1 No. 1 dan Pasal 1 No. 4 UU No. 11 Tahun 2008 jo. Undang-Undang Nomor 19 Tahun 2016 tentang Informasi dan Transaksi Elektronik.

Informasi elektronik adalah satu atau lebih dari satu klaster data elektronik, termasuk namun tidak terbatas pada tulisan, suara, gambar, peta, rancangan, foto, Pertukaran data elektronik, surat elektronik, telegram, teleks, telekopi atau sejenisnya, huruf, tanda, angka, Kode Akses, simbol atau perforasi yang telah diolah maknanya atau dimengerti oleh orang yang memenuhi syarat untuk memahaminya.

Dokumen Elektronik adalah Informasi Elektronik yang dibuat, diteruskan, dikirim, diterima, atau disimpan dalam bentuk analog, digital, elektromagnetik, optik, atau sejenisnya, terlihat, dapat ditampilkan, dan/atau terdengar melalui Komputer atau Sistem Elektronik, termasuk namun tidak terbatas pada tulisan, suara, gambar, peta, rancangan, foto atau sejenisnya, huruf, tanda, angka, Kode Akses, simbol, atau perforasi yang memiliki arti atau definisi tertentu atau dapat dipahami oleh orang yang memenuhi syarat untuk memahaminya.

Susunan kata yang mirip dapat ditemukan dalam Undang-undang sektoral, seperti Undang-Undang Keterbukaan Informasi Publik (UU No. 14 Tahun 2008), yang mendefinisikan informasi sebagai “pernyataan, gagasan, dan tanda yang memiliki nilai, makna, dan pesan, baik data, fakta, dan klarifikasi yang dapat dilihat, didengar, dan dibaca, serta disajikan dalam berbagai paket dan format, sesuai dengan perkembangan teknologi informasi dan komunikasi, baik secara elektronik maupun non-elektronik”.⁸ Maka dari itu, informasi elektronik dan dokumen elektronik apapun yang berpotensi mengandung substansi ilegal dan berbahaya dapat dikategorisasikan sebagai konten “ilegal” dan “berbahaya”.

⁸ Pasal 11 No. 1 Undang-Undang Nomor 14 Tahun 2008 tentang Keterbukaan Informasi Publik.

Pada dasarnya, UU ITE sebagai landasan hukum utama dalam mengatur konten tidak membedakan antara konten ilegal dan berbahaya. Namun, Penjelasan Umum UU ITE 2016 menggunakan istilah “konten ilegal” untuk merujuk pada informasi elektronik dan/atau dokumen elektronik yang memuat:⁹

- Konten yang melanggar kesucilaan;
- Konten perjudian;
- Fitnah atau pencemaran nama baik;
- Pemerasan dan/atau ancaman;
- Berita bohong dan menyesatkan yang mengakibatkan kerugian konsumen dalam transaksi elektronik;
- kebencian atau permusuhan berdasarkan suku, agama, ras, dan golongan; dan
- ancaman kekerasan atau intimidasi yang ditujukan kepada individu.

Konten ilegal ini secara substansial diatur dalam Bab VII tentang Perbuatan yang Dilarang dari UU ITE, khususnya dari Pasal 27 sampai dengan Pasal 29. Pasal-pasal dalam Bab ini adalah ketentuan mengenai Judul 1 dan Judul 2 Kejahatan Siber¹⁰ seperti *hacking, cracking, dan phishing*.¹¹

Maka dari itu, meskipun beberapa instrumen hukum internasional dan regional di atas menyoroti pentingnya membedakan mana yang ilegal dan mana yang berbahaya, dalam konteks UU ITE, konten yang berpotensi “berbahaya” yang disebutkan dalam beberapa instrumen hukum internasional atau regional di atas pada dasarnya diklasifikasikan sebagai konten ilegal menurut UU ITE. Sebagai konsekuensinya, karena perbuatan

⁹ Lihat Penjelasan Umum Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.

¹⁰ Lihat Pasal 2 sampai Pasal 8 the Convention of Cybercrime, Budapest, 23.XI.2001.

¹¹ Lihat Pasal 30 sampai dengan Pasal 37 Undang-Undang Nomor 11 Tahun 2008 jo. Undang-Undang Nomor 19 Tahun 2016 tentang Informasi dan Transaksi Elektronik.

yang dilarang itu dianggap sebagai tindak pidana, maka pelakunya dapat dituntut secara pidana.

Meskipun UU ITE telah menjadi dasar hukum utama untuk mengatur konten daring, namun UU tersebut tidak dapat berdiri sendiri. Menurut Mahkamah Konstitusi Republik Indonesia, UU ITE pada dasarnya memperluas cakupan delik dari ruang fisik ke ruang siber.¹² Oleh karena itu, kita juga harus merujuk ke tindakan lain untuk menentukan apakah konten daring dianggap ilegal/berbahaya dalam beberapa kasus.

Ruang lingkup konten ilegal yang disebutkan di atas diperluas lebih lanjut dalam Peraturan Pemerintah (PP) tentang Penyelenggaraan Sistem dan Transaksi Elektronik sebagai peraturan pelaksana utama dari UU ITE.¹³ Peraturan Pemerintah ini secara implisit “membedakan” jenis konten menjadi dua kategori, yaitu (a) konten yang melanggar peraturan perundang-undangan; dan (b) konten yang mengganggu masyarakat dan ketertiban umum.

Klasifikasi pertama meliputi yang memuat “muatan yang dilarang sesuai dengan ketentuan peraturan perundang-undangan”, yang pada hakikatnya berarti “konten ilegal”. Kategori pertama meliputi informasi elektronik dan dokumen elektronik yang mengandung unsur (a) pornografi, (b) perjudian, (c) fitnah, (d) pencemaran nama baik, (e) penipuan, (f) kebencian terhadap suatu komunitas berdasarkan suku, agama, ras, dan antar golongan, (g) kekerasan dan/atau kekerasan terhadap anak, (h) pelanggaran kekayaan intelektual, (i) pelanggaran perdagangan barang dan jasa melalui sistem elektronik, (j) terorisme dan/atau radikalisme, separatisme dan/atau organisasi berbahaya yang dilarang, (k) pelanggaran keamanan informasi, (l) pelanggaran perlindungan konsumen, (m) pelanggaran di bidang kesehatan, (n) pelanggaran pengawasan obat dan makanan.¹⁴

¹² Lihat misalnya Putusan Mahkamah Konstitusi Nomor 50/PUU-VI/2008 tentang Pengujian Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik Terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945.

¹³ Lihat Pasal 95 dan Pasal 96 Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik.

¹⁴ Lihat Pasal 96 huruf a Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik.

Sebagian besar konten tersebut di atas, secara substansial, diatur sebagai perbuatan yang dilarang dalam UU ITE, termasuk pornografi, perjudian, fitnah, pencemaran nama baik, penipuan, dan kebencian terhadap masyarakat, berdasarkan suku, agama, ras, dan antar golongan. Bahkan konten online yang mengandung kekerasan dan/atau kekerasan terhadap anak, terorisme dan/atau radikalisme, separatisme, dan/atau organisasi berbahaya yang dilarang dapat merupakan tindakan terlarang dalam beberapa kasus. Jenis konten lain juga harus mengacu pada Undang-undang lain, seperti KUHP, Undang-Undang Hak Cipta, Undang-Undang Kesehatan, Undang-Undang Pangan, dan Undang-Undang Perlindungan Konsumen.

Kategori kedua mengenai konten yang mengganggu ketertiban umum dan masyarakat dapat diartikan sebagai “berbahaya”, walau tampaknya mengarah ke konten yang tidak melanggar peraturan perundang-undangan tetapi tetap bersifat merugikan bagi masyarakat dan ketertiban umum. Konten terlarang yang disebutkan di dalam Peraturan Pemerintah (PP) tentang Penyelenggaraan Sistem dan Transaksi Elektronik tidak dimaksudkan untuk memberikan sanksi kepada individu (seperti dalam UU ITE) tetapi untuk mengontrol konten yang disediakan atau disebarluaskan melalui sistem elektronik, yang disediakan oleh operator sistem elektronik.

Meskipun kerangka hukum saat ini menyebutkan berbagai jenis konten yang dapat diklasifikasikan sebagai “ilegal” atau “berbahaya”, penelitian ini hanya akan mengkaji ketentuan yang berkaitan dengan pencemaran nama baik, ujaran kebencian, dan berita bohong. Terlepas dari cakupan proyek SM4P secara keseluruhan, pemilihan konten juga didasarkan pada beberapa laporan dari organisasi masyarakat sipil di Indonesia, yang menunjukkan bahwa ketentuan terkait pencemaran nama baik, ujaran kebencian, dan berita bohong adalah salah satu dasar hukum yang paling banyak digunakan untuk penuntutan pidana atas ucapan seseorang di media sosial (Lihat, misalnya, Tempo, 2020; dan SAFEnet, 2021). Selain ketiga jenis konten di atas, penelitian ini juga menyoroti satu lagi klasifikasi konten yang mengkritik dan berpotensi membatasi kebebasan berekspresi

di media sosial, yaitu konten yang mengganggu masyarakat dan ketertiban umum. Elaborasi disediakan di bawah ini.

➔ **Fitnah dan/atau Pencemaran Nama Baik**

Ketentuan tentang fitnah dan pencemaran nama baik dalam UU ITE dianggap sebagai salah satu ketentuan yang paling banyak digunakan (Lihat contohnya di Tempo, 2020; dan SAFEnet, 2021). Ketentuan ini juga merupakan ketentuan UU ITE pertama yang diajukan untuk pengujian konstitusional.¹⁵ Selanjutnya fitnah dan pencemaran nama baik dalam UU ITE juga tidak terlepas dari fitnah dan pencemaran nama baik yang diatur dalam KUHP.¹⁶ KUHP mendefinisikan tindak pidana fitnah dan/atau pencemaran nama baik sebagai: “[...] dengan sengaja menyerang kehormatan atau nama baik seseorang dengan menuduh seseorang melakukan sesuatu, dengan maksud yang jelas untuk memberikan publisitas.”¹⁷

Interpretasi ini ditegaskan oleh Mahkamah Konstitusi dalam Perkara No. 50/PUU-VI/2008, dimana Mahkamah menyatakan bahwa interpretasi dari pencemaran nama baik atau fitnah dalam Pasal 27 ayat (3) UU ITE tidak terlepas dari norma hukum pidana yang termaktub dalam Pasal 310 dan Pasal 311 KUHP.¹⁸ Pengadilan berpendapat bahwa:¹⁹

¹⁵ Putusan Mahkamah Konstitusi Nomor 50/PUU-VI/2008 tentang Pengujian Undang-Undang Nomor 11 Tahun 2008 Tentang Informasi dan Transaksi Elektronik Terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (4 Mei 2009).

¹⁶ Lihat misalnya, Putusan Pengadilan Negeri Bantul No. 125/Pid.Sus/2018/PN Btl (30 Januari 2019) 22-23; Putusan Pengadilan Negeri Jantoro No. 76/Pid.Sus/2021/PN Jth (27 September 2021) 19.

¹⁷ Lihat Pasal 310 ayat (1) KUHP.

¹⁸ Lihat Putusan Mahkamah Konstitusi Nomor 50/PUU-VI/2008 tentang Pengujian Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik Terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (4 Mei 2009) 110.

¹⁹ *Ibid.*

[3.17.1] Bahwa terlepas dari pertimbangan Mahkamah yang telah diuraikan dalam paragraf terdahulu, keberlakuan dan tafsir atas Pasal 27 ayat (3) UU ITE tidak dapat dipisahkan dari norma hukum pokok dalam Pasal 310 dan Pasal 311 KUHP sebagai genus delict yang mensyaratkan adanya pengaduan (klacht) untuk dapat dituntut, harus juga diperlakukan terhadap perbuatan yang dilarang dalam Pasal 27 ayat (3) UU ITE, sehingga Pasal a quo juga harus ditafsirkan sebagai delik yang mensyaratkan pengaduan (klacht) untuk dapat dituntut di depan Pengadilan.

Selain itu, baik UU ITE maupun KUHP tidak memberikan penjelasan yang cukup tentang apa yang merupakan unsur suatu perbuatan yang mengandung konten 'penghinaan dan/atau pencemaran nama baik'. Misalnya, dalam perkara No. 132/PID.B/2010/PN.MRK, hakim hanya menginterpretasi unsur 'sengaja menyerang kehormatan atau reputasi seseorang dengan menuduh seseorang melakukan sesuatu' sebagai niat terdakwa untuk menyerang kehormatan atau reputasi seseorang dengan menuduh seseorang melakukan sesuatu.²⁰ Mahkamah Konstitusi, dalam putusan tersebut di atas, juga menegaskan kembali bahwa harus ada perbedaan antara "dengan sengaja melakukan suatu perbuatan" dan "dengan sengaja menyerang kehormatan atau nama baik orang lain" dalam Pasal 310.²¹

Selanjutnya, dengan tidak adanya definisi atau interpretasi yang jelas tentang konten pencemaran nama baik, beberapa pengadilan memperluas objek pencemaran nama baik, termasuk kehormatan badan hukum atau lembaga negara.²² Misalnya, dalam Perkara No. 223/Pid.Sus/2018/PN Kbm, Mahkamah mengacu pada Putusan Mahkamah Agung No. 183 K/Pid/2010

²⁰ Lihat Putusan Pengadilan Negeri Merauke No. 132/PID.B/2010/PN.MRK, hal. 13, Ad. 2.

²¹ Lihat Putusan Mahkamah Konstitusi No. 50/PUU-VI/2008, hal. 42.

²² Lihat misalnya, Putusan Pengadilan Negeri Kebumen No. 223/Pid.Sus/2018/PN Kbm (17 Desember 2018) 41.

yang menyatakan bahwa badan hukum dapat menjadi objek pencemaran nama baik.²³

Dengan berbagai polemik yang terjadi akibat perbedaan interpretasi dari pihak polisi, jaksa, dan pengadilan mengenai perbuatan yang dilarang dalam UU ITE, Menteri Komunikasi dan Informatika (Menkominfo), Kejaksaan Agung, dan Kapolri mengeluarkan pedoman pelaksanaan UU ITE (Keputusan Bersama).²⁴ Khusus untuk Pasal 27 ayat (3), Keputusan Bersama tersebut menegaskan bahwa fokus pemidanaan dalam Pasal ini bukan pada perasaan korban tetapi perbuatan pelaku.²⁵ Lebih lanjut, Keputusan Bersama ini menyatakan bahwa korban sebagai pelapor harus merupakan individu dengan identitas tertentu dan bukan lembaga, korporasi, profesi, atau jabatan.²⁶ Sepintas, hal ini tampak bertentangan dengan Putusan Mahkamah Agung No. 183 K/Pid/2010 atas. Meski demikian, Keputusan Bersama ini tidak menspesifikkan apakah suatu lembaga dapat dijadikan objek pencemaran nama baik atau fitnah. Dengan demikian, masih dapat diartikan bahwa suatu lembaga bisa menjadi objek pencemaran nama baik atau fitnah. Namun, pelapor harus merupakan individu tertentu. Perumusan konten fitnah/pencemaran nama baik dalam UU ITE telah mempengaruhi pelaksanaan kebebasan berekspresi di Indonesia dan mendapat banyak kritikan dari berbagai organisasi masyarakat sipil.²⁷

➔ **Berita Bohong dan Menyesatkan (Misinformasi/Disinformasi)**

● **Berita Bohong dan Menyesatkan Secara Umum (KUHP)**

Misinformasi/Disinformasi secara umum diatur dalam KUHP. Pasal-pasal yang mengatur tentang larangan menyebarkan berita atau informasi bohong antara lain Pasal 14 ayat (1), 14 ayat (2), dan Pasal 15

²³ Lihat Putusan MA No. 183 K/Pid/2010 (20 Mei 2010) 15.

²⁴ Keputusan Bersama Menteri Komunikasi dan Informatika, Kejaksaan Agung, dan Kapolri Nomor 229 Tahun 2021, Nomor 154 Tahun 2021, Nomor KB/2/VI/2021.

²⁵ *Ibid* 12.

²⁶ *Ibid*.

²⁷ Penjelasan lebih lanjut mengenai kekhawatiran, tren, dan dampak dari undang-undang pencemaran nama baik akan dijelaskan pada Bab II dan Bab III.

KUHP. Namun, pasal-pasal tersebut hanya menetapkan istilah 'berita atau informasi bohong yang dapat menerbitkan keonaran dikalangan rakyat'. Pasal-pasal ini tidak menjelaskan lebih lanjut jenis-jenis berita atau informasi apa yang dikategorikan sebagai "berita atau informasi bohong, dan pada tingkat apa masalah ini dapat menimbulkan keonaran dikalangan rakyat" yang dimaksud dalam pasal-pasal tersebut. Dengan tidak adanya ketentuan tersebut, penting bagi kita untuk melihat yurisprudensi tentang pasal-pasal ini.

Karena cakupannya yang luas dibandingkan dengan pasal-pasal misinformasi/disinformasi dalam UU ITE, pasal-pasal tersebut juga digunakan untuk kasus-kasus di dunia maya, terutama terkait dengan misinformasi. Misalnya, dalam Perkara No. 203/Pid.Sus/2019/PN.Jkt.Sel tentang kasus misinformasi di media sosial, jaksa menggunakan dakwaan alternatif dengan menggunakan Pasal 14(1) KUHP dan Pasal 28(2) UU ITE tentang ujaran kebencian.

Setelahnya, Mahkamah berpendapat bahwa terdakwa terbukti melakukan tindak pidana sebagaimana dimaksud dalam dakwaan pertama. Namun, Pengadilan tidak secara eksplisit mendefinisikan apa yang dimaksud dengan "berita bohong", dan alih-alih mengaitkannya dengan fakta-fakta yang diberikan oleh para saksi. MK lebih mempermasalahkan dampak yang terjadi karena "berita" tersebut menjadi viral dan menjadi trending topic. Menurut Mahkamah, viralitas ini merupakan "benih" kekacauan yang muncul ke permukaan.²⁸

Sementara itu, Perkara No. 471/Pid.Sus/2020/PN.Bdg mendefinisikan "berita bohong" sebagai informasi yang dibuat-buat atau disampaikan secara tidak benar--tidak sesuai dengan fakta dan sengaja dibuat untuk menyesatkan penerima informasi.²⁹ Pengadilan menekankan bahwa niat/tujuan dari publikasi sebuah "berita bohong"

²⁸ Lihat Putusan Pengadilan Negeri Jakarta Selatan No. 203/Pid.Sus/2019/PN.Jkt.Sel, hal. 142, paragraf. 7.

²⁹ Lihat Putusan Pengadilan Negeri Bandung No. 471/Pid.Sus/2020/PN.Bdg, hal. 121, Ad. 2.

sangat penting, karena "berita bohong" diterbitkan untuk membuat orang percaya bahwa konten tersebut asli.³⁰ Dari kedua putusan tersebut terlihat adanya perbedaan penafsiran atas istilah-istilah yang terdapat dalam pasal 14(1), 14(2), dan 15. Ketiadaan definisi dan tingkatan yang jelas tersebut dapat menyebabkan penerapan aturan yang sewenang-wenang oleh pihak berwajib.

Lebih lanjut, menurut PP PSTE, konten misinformasi/disinformasi juga tergolong konten yang mengganggu masyarakat dan ketertiban umum.³¹ Meski tergolong dalam kategori yang berbeda dengan konten yang melanggar peraturan perundang-undangan, berdasarkan penjelasan di atas, konten disinformasi tetap tergolong konten yang melanggar KUHP. Dalam beberapa kasus, penggolongan konten yang melanggar KUHP ini juga dapat digunakan secara setara/ sepadan dengan ketentuan tentang tindakan yang dilarang dalam UU ITE.

Berita Bohong dan Menyesatkan yang Mengakibatkan Kerugian Konsumen dalam Transaksi Elektronik (UU ITE)

Berita bohong dan menyesatkan dalam Pasal 28 ayat (1) UU ITE berbeda dengan misinformasi/disinformasi dalam KUHP.³² Unsur-unsur yang diatur dalam UU ITE lebih dipersempit dibandingkan dengan KUHP, karena hanya mencakup informasi bohong dan menyesatkan dalam konteks transaksi elektronik. Pasal 28 ayat (1) UU ITE bertujuan untuk melindungi konsumen dalam transaksi elektronik. Menurut UU ITE, pengertian transaksi elektronik adalah "tindakan legal yang dilakukan dengan menggunakan komputer, jaringan komputer, dan/atau media elektronik lainnya."³³ Oleh karena itu, seseorang yang hanya mendistribusikan atau menyebarkan disinformasi atau berita bohong tidak cukup untuk diklasifikasikan sebagai tindakan yang dilarang berdasarkan pasal ini. Unsur

³⁰ *Ibid.*, p.130, Ad.3.

³¹ Lihat Pasal 96 huruf b dan Penjelasan Pasal 96 huruf b Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik.

³² *Ibid.*, 16.

³³ Lihat Pasal 1 No.3 Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.

“mengakibatkan kerugian konsumen” juga perlu dibuktikan.³⁴

Konten dalam pasal ini juga dapat termasuk dalam kategori konten yang mengandung pelanggaran perdagangan barang dan jasa melalui sistem elektronik. Selain itu, pengertian konsumen dalam konteks Pasal 28 ayat (1) UU ITE harus mengacu pada Undang-Undang Nomor 8 Tahun 1999 tentang Perlindungan Konsumen. Undang-undang tersebut mendefinisikan konsumen sebagai “pengguna dari barang dan/atau jasa yang tersedia di masyarakat, baik untuk kepentingan dirinya sendiri, kerabat, orang lain, maupun makhluk hidup lainnya, dan bukan untuk diperdagangkan”.³⁵ Terlebih lagi, berita bohong dan menyesatkan dalam UU ITE juga bisa terkait dengan Pasal 378 KUHP yang mengatur tentang penipuan.

➔ **Kebencian atau Permusuhan Berdasarkan Suku, Agama, Ras, dan Antargolongan (Ujaran Kebencian)**

Pasal tentang ujaran kebencian dalam Pasal 28 ayat (2) UU ITE juga merupakan salah satu Pasal yang paling banyak digunakan dalam UU ITE (Lihat contohnya, dalam Tempo, 2020; dan SAFEnet, 2021). Pasal tentang ujaran kebencian dapat dikaitkan dengan beberapa Undang-undang, antara lain KUHP dan UU Penghapusan Diskriminasi Ras dan Etnis serta UU Pencegahan Penyalahgunaan Agama dan/atau Penodaan Agama.

Dalam KUHP Indonesia, ada beberapa pasal yang mengatur larangan mengungkapkan perasaan permusuhan, kebencian, atau penghinaan, yaitu Pasal 156, Pasal 156a, dan Pasal 157. Pasal 156 KUHP sebenarnya adalah 'induk' dari Pasal 156a. Pasal 156 KUHP mengatur perbuatan terhadap 'sesuatu atau beberapa golongan penduduk Negara Indonesia'. Sedangkan Pasal 156a mengatur tentang tindakan terhadap salah satu golongan tersebut, yaitu golongan agama (Lembaga Peradilan Independen Indonesia, 2018).

³⁴ Lihat misalnya Putusan Pengadilan Negeri Jakarta Timur No. 532/Pid.Sus/2020/PN.Jkt.Tim (6 Oktober 2020) 21-22; Putusan Pengadilan Negeri Bale Bandung No. 84/Pid.Sus/2021/PN.Bib (19 April 2021) 17-19.

³⁵ Pasal 1 No. 2 Undang-Undang Nomor 8 Tahun 1999 tentang Perlindungan Konsumen.

Terminologi terkait dapat merujuk pada UU Penghapusan Diskriminasi Ras dan Etnis. Undang-undang ini menerangkan banyak bentuk yang dapat dikategorikan sebagai “diskriminasi ras dan etnis”, antara lain membuat tulisan atau gambar untuk ditempatkan, ditempel, atau disebarluaskan di tempat umum atau tempat lain yang dapat dilihat atau dibaca orang lain, memberikan pidato, mengungkapkan, atau mengucapkan kata-kata tertentu di tempat umum atau tempat lain yang dapat didengar oleh orang lain, mengenakan sesuatu pada dirinya berupa benda, kata, atau gambar di tempat umum atau tempat lain yang dapat dibaca orang lain, melakukan perampasan nyawa orang, penyiksaan, pemerkosaan, perbuatan cabul, pencurian dengan kekerasan, atau perampasan kemerdekaan berdasarkan diskriminasi ras dan etnis.

Beberapa diskursus mengenai pasal-pasal ujaran kebencian melibatkan istilah-istilah seperti “di depan umum”, “dengan maksud agar isinya diketahui atau lebih dikenal publik”, dan “antar kelompok”. Misalnya, dalam konteks terminologi “di depan umum”, hakim dalam perkara No. 1537/Pid.B/2016/PN JKT.UTR menafsirkan “di depan umum” dengan merujuk pada buku “Delik - Delik Khusus Kejahatan Terhadap Kepentingan Hukum” oleh Drs. PAF Lamintang, SH yang menyatakan bahwa “di muka umum” dalam rumusan pidana yang diatur dalam Pasal 156a KUHP tidak berarti bahwa perasaan yang diungkapkan oleh pelaku atau perbuatan yang dilakukan oleh pelaku harus selalu terjadi di tempat umum, tetapi sudah cukup apabila emosi yang diungkapkan oleh pelaku dapat didengar oleh publik, atau tindakan yang dilakukan oleh pelaku dapat dilihat oleh publik.³⁶ Penjelasan Undang-Undang Pencegahan Penyalahgunaan dan/atau Penodaan Agama juga merujuk istilah “di depan umum” pada terminologi yang digunakan dalam KUHP.

Mengenai terma/istilah kedua, niat untuk membuat konten menjadi lebih dikenal publik dapat diketahui melalui sikap batin terhadap kesadaran bahwa tindakan mereka memang dapat diketahui oleh publik (Purwati, 2018, hlm. 43). Sikap batin seseorang tentu bukan sesuatu yang mudah

³⁶ Lihat Putusan Pengadilan Negeri Jakarta Utara No.1537/Pid.B/2016/PN JKT.UTR, hal. 594.

dibuktikan. Oleh karena itu, niat tersebut dianggap tercermin ketika pelaku melakukan tindakannya di depan umum. Pada dasarnya, terdapat maksud di sini untuk mempertegas bagian “di depan umum” yang terdapat dalam Pasal 157.

Mengenai istilah “antar golongan”, meskipun kita dapat merujuk pada beberapa Undang-undang yang telah disebut di atas, dalam praktiknya istilah ini dapat diartikan secara luas oleh para polisi, jaksa, dan pengadilan. Mahkamah Konstitusi juga mengakui penafsiran yang luas tentang “antar kelompok”. Dalam Perkara No. 76/PUU-XV/2017, Mahkamah memutuskan bahwa:³⁷

[3.13.1] Bahwa istilah “antargolongan”, menurut Mahkamah, tidak jelas dan tegas. Makna istilah tersebut tidak serta merta dapat diketahui, berbeda dengan istilah “suku”, “agama”, dan “ras”, yang ditempatkan sejajar dengan istilah “antargolongan” bahkan memunculkan singkatan yang populer di masyarakat, yaitu SARA. Meski tidak jelas dan tegas, bukan berarti “antargolongan” itu tidak ada. [...]

Yang dimaksud dengan “antargolongan” jelas bukan “antargolongan” sebagaimana dimaksud dalam Pasal 163 dan Pasal 131 IS, melainkan pada realitas sosiologis tentang adanya “kelompok lain” di luar suku, agama, dan ras. [...]

Selain itu, juga rumit untuk membedakan konten mana yang menyinggung dan dimaksudkan untuk menimbulkan permusuhan dan kebencian. Dalam Perkara No. 366/Pid.Sus/2019/PN.JKT.SEL, misalnya, Putusan Pengadilan Negeri Jakarta Selatan menyatakan bahwa konten yang menyinggung dapat dikategorikan sebagai penyebab kebencian dan permusuhan.³⁸ Rinciannya dinyatakan sebagai berikut:

³⁷ Putusan Mahkamah Konstitusi Nomor 76/PUU-XV/2017 tentang Pengujian Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana telah diubah dengan Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (27 Maret 2018) 66-68.

³⁸ Lihat misalnya, Putusan Pengadilan Negeri Jakarta Selatan No. 366/Pid.Sus/2019/PN.JKT.SEL (15 Agustus 2019) 84-85.

[...] Menimbang, bahwa berdasarkan pertimbangan-pertimbangan, apa yang dibuktikan dalam perkara ini adalah kesengajaan dengan suatu kepastian yang disadari, seperti dalam perkara ini, Tergugat memahami atau menyadari bahwa perbuatannya, akibat-akibatnya, dan keadaan yang menyertainya, bahwa kedudukan Tergugat akan menyinggung etnis Tionghoa dan golongan pendukung Capres dan Cawapres Nomor 1, dan Calon Presiden dan Wakil Presiden Nomor 1 Joko Widodo dan Ma'aruf Amin, serta KPU selaku penyelenggara pemilihan umum 2019.

Selain itu, rincian diatas juga menggambarkan bahwa MK memaknai secara luas istilah “kelompok”, yang juga mencakup kelompok orang (Capres dan Wakil Presiden) dan lembaga negara (KPU). Interpretasi yang lebih luas dari “kelompok” dapat ditemukan dalam Perkara No. 16/Pid.Sus/2020/PN Jap, di mana Pengadilan Negeri Jayapura menyatakan bahwa Bangsa Indonesia (Bangsa Indonesia) juga merupakan “kelompok”.³⁹ Mahkamah menyatakan bahwa unggahan yang dibuat oleh Terdakwa menimbulkan permusuhan, dalam hal ini adalah permusuhan terhadap bangsa Indonesia.

Dalam kasus lain, Pengadilan Negeri Kendari menyatakan bahwa kritik keras dan komentar negatif terhadap lembaga negara juga bisa masuk dalam Pasal 28 ayat (2) UU ITE. Mahkamah berpendapat bahwa apabila seseorang memahami akan adanya berbagai tanggapan, baik yang pro maupun kontra, yang menimbulkan polemik sosial melalui media sosial, maka konten yang mereka posting dapat dikategorikan sebagai konten yang menimbulkan kebencian atau permusuhan antarkelompok masyarakat.⁴⁰

³⁹ Lihat Putusan PN Jayapura No. 16/Pid.Sus/2020/PN Jap (29 April 2020) 38-39.

⁴⁰ Lihat Putusan Pengadilan Negeri Kendari No. 426/Pid.Sus/2021/PN Kdi (16 September 2021) 24. Lihat juga Keputusan Bersama Menteri Komunikasi dan Informatika, Kejaksaan Agung, dan Kapolri No. 229 Tahun 2021, No. 154 Tahun 2021, No. KB/2/VI/2021, 20-21.

Pasal ini juga dapat digunakan untuk menyasar publikasi atau penyebaran konten yang mengandung terorisme dan/atau radikalisme, separatisme, dan/atau organisasi berbahaya yang dilarang.

Tidak adanya batasan yang jelas tentang bagaimana konten dapat dianggap sebagai kebencian atau permusuhan memiliki potensi tinggi untuk melanggar kebebasan berekspresi warga negara yang dilindungi oleh UUD 1945. Hal ini dapat dilihat dari berbagai penafsiran yang dilakukan oleh beberapa pengadilan negeri sebagaimana tersebut di atas, yang dapat memperluas makna pasal-pasal tersebut.

Selanjutnya dalam konteks publikasi melalui pers, UU Pers dapat dijadikan sebagai salah satu landasan hukum, karena UU ini melarang iklan menghasilkan kerugian suatu agama dan/atau mengganggu kerukunan antar umat beragama, atau yang bertentangan dengan rasa kesucilaan masyarakat. Namun, dalam UU Pers tidak ada definisi yang tegas mengenai apa yang dimaksud dengan “merendahkan suatu agama dan/atau mengganggu kerukunan antarumat beragama, dan bertentangan dengan rasa kesucilaan masyarakat”. Dengan demikian, penafsiran istilah-istilah tersebut dibiarkan terbuka lebar dan rawan dinilai secara subjektif.

→ Contents that Disturb Community and Public Order

Meski terkesan berbeda dengan konten yang melanggar peraturan perundang-undangan menurut Pasal 96 huruf a PP PSTE, namun konten jenis ini masih bisa diartikan secara luas. Sampai saat ini, belum ada penerangan yang jelas tentang apa yang mengganggu masyarakat dan ketertiban umum pada tingkat UU. Akibatnya, konten ini bisa dimaknai secara luas oleh lembaga negara dan penegak hukum. Dalam beberapa peraturan daerah, ketertiban umum diartikan sebagai “keadaan dimana pemerintah dan warga negara dapat melaksanakan kegiatannya dengan tertib”.⁴¹ Dalam konteks Peraturan Daerah tentang ketertiban umum, ruang lingkup ketertiban umum dapat berbeda-beda menurut Daerah.

⁴¹ Lihat misalnya Peraturan Daerah Daerah Khusus Ibukota Jakarta Nomor 8 Tahun 2007 tentang Ketertiban Umum; Peraturan Daerah Daerah Istimewa Yogyakarta Nomor 2 Tahun 2017 tentang Ketentraman, Ketertiban Umum, dan Perlindungan Masyarakat; dan Peraturan Daerah Kabupaten Buton Nomor 2 Tahun 2020 tentang Penyelenggaraan Ketentraman, Kedamaian, Ketertiban Umum, dan Perlindungan Masyarakat.

Penjelasan Pasal 96 huruf b menjelaskan bahwa yang dimaksud dengan “menggangu masyarakat dan ketertiban umum” antara lain meliputi informasi dan/atau fakta yang dipalsukan.⁴² Meskipun publikasi informasi yang dipalsukan tidak diperinci sebagai tindakan yang dilarang dalam UU ITE, namun ia tetap dapat dikaitkan dengan tindakan terlarang lainnya berdasarkan UU ITE, seperti kebencian dan/atau permusuhan terhadap komunitas berdasarkan suku, agama, ras, dan antar golongan,⁴³ dan juga termasuk dalam ruang lingkup Pasal 14 KUHP. Meskipun informasi yang dipalsukan tidak tergolong konten yang melanggar peraturan perundang-undangan di bawah PP PSTE, namun tetap dapat dituntut berdasarkan KUHP dan UU ITE. Oleh karena itu, contoh yang diberikan dalam penjelasan Pasal 96 huruf b pada dasarnya merupakan konten yang “melanggar peraturan perundang-undangan” juga.

Mengingat Pasal 96 huruf a dan huruf b pada prinsipnya sama, maka pembedaan klasifikasi konten yang “melanggar peraturan perundang-undangan” dan “menggangu ketertiban umum” menjadi dipertanyakan. Akan muncul pertanyaan lebih lanjut mengenai jenis konten lain yang menggangu masyarakat dan ketertiban umum, karena penjelasannya hanya menyebutkan satu jenis konten. Terminologi yang digunakan terlalu kabur dan dapat diartikan secara luas sesuai dengan kepentingan pemerintah atau negara. Penjelasan tersebut hanya memberikan satu contoh dari apa yang dapat dikategorikan sebagai “menggangu masyarakat dan ketertiban umum”. Oleh karena itu, potensi pelanggaran HAM berdasarkan ketentuan ini cukup tinggi, terutama mengenai kebebasan berekspresi di media sosial. Jika suatu konten dianggap menggangu ketertiban umum dan masyarakat, bisa saja konten tersebut diblokir oleh platform media sosial berdasarkan permintaan dari pemerintah.

Mengingat Pasal 96 poin a dan huruf b pada prinsipnya sama, maka pembedaan klasifikasi konten yang “melanggar peraturan perundang-

⁴² Lihat Pasal 96 huruf b Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik.

⁴³ Lihat misalnya, Putusan Pengadilan Negeri Jakarta Selatan No. 366/Pid.Sus/2019/PN.JKT.SEL (15 Agustus 2019) 77-78.

undangan” dan “mengganggu ketertiban umum” menjadi dipertanyakan. Akan muncul pertanyaan lebih lanjut mengenai jenis konten lain yang mengganggu masyarakat dan ketertiban umum, karena penjelasannya hanya menyebutkan satu jenis konten. Terminologi yang digunakan tidak terlalu jelas dan dapat diartikan secara luas sesuai dengan kepentingan pemerintah atau negara. Penjelasan tersebut hanya memberikan satu contoh dari apa yang dapat dikategorikan sebagai “mengganggu masyarakat dan ketertiban umum”. Oleh karena itu, potensi pelanggaran HAM berdasarkan ketentuan ini cukup tinggi, terutama mengenai kebebasan berekspresi di media sosial. Jika suatu konten dianggap mengganggu ketertiban umum dan masyarakat, bisa saja konten tersebut diblokir oleh platform media sosial berdasarkan permintaan dari pemerintah.

➔ **Konsekuensi Regulasi Konten Online**

Seperti yang disebutkan di atas, menurut UU ITE, “konten online” dapat merujuk pada informasi dan dokumen elektronik apa pun. Namun, polisi juga perlu merujuk undang-undang dan peraturan lain dalam menangani konten ilegal dan berbahaya secara online. Tabel di bawah ini merangkum jenis konten online yang diatur dalam beberapa UU, termasuk lima kategori konten yang dijelaskan di atas.

Tabel 1. Klasifikasi Konten Daring Berdasarkan Peraturan Indonesia

Tipe Konten	Klasifikasi	Regulasi Terkait ⁴⁴
konten yang melanggar kepatutan (termasuk pornografi dan pornografi	Pidana	<ul style="list-style-type: none"> ● UU ITE 2008 jo. 2016 ● KUHP 1946 ● UU Pornografi 2008 ● PP PSTE 2019

⁴⁴ Peraturan-peraturan yang dicantumkan mulai dari dasar hukum utama/paling relevan berdasarkan hierarki peraturan, dan dilanjutkan dengan beberapa peraturan terkait. Selain PP PSTEI, peraturan pelaksanaan lainnya dari undang-undang tertentu tidak dicantumkan secara eksplisit tetapi masih dapat digunakan sebagai dasar teknis pelaksanaan undang-undang tersebut.

konten tentang judi	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • KUHP 1946 • PP PSTE 2019
konten tentang penghinaan dan/atau pencemaran nama baik	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • KUHP 1946 • PP PSTE 2019
konten tentang pemerasan dan/atau ancaman	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • KUHP 1946 • PP PSTE 2019
informasi salah dan menyesatkan yang mengakibatkan kerugian konsumen dalam Transaksi Elektronik	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • UU Perlindungan Konsumen 1999 • PP PSTE 2019
informasi yang ditujukan untuk menimbulkan kebencian atau pertikaian terhadap individu dan/atau kelompok masyarakat tertentu berdasarkan suku, agama, ras, dan antargolongan	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • KUHP 1946 • UU Penghapusan Diskriminasi Ras dan Etnis 2008 • UU Pencegahan Penyalahgunaan dan/atau Penodaan Agama 1965 • PP PSTE 2019

mengandung ancaman kekerasan atau menakutkan terhadap seorang individu	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • KUHP 1946 • PP PSTE 2019
konten yang mengandung elemen kekerasan terhadap anak	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • UU Perlindungan Anak 2002 jo. 2014 • PP PSTE 2019
konten yang mengandung elemen pelanggaran hak cipta	Pidana	<ul style="list-style-type: none"> • UU Hak Cipta 2014 • UU ITE 2008 jo. 2016 • PP PSTE 2019
konten yang mengandung elemen terorisme dan/atau radikalisme, separatisme dan/atau organisasi berbahaya terlarang	Pidana	<ul style="list-style-type: none"> • UU Terorisme 2003 jo. 2018 • UU ITE 2008 jo. 2016 • PP PSTE 2019
konten yang mengandung elemen pelanggaran dalam perdagangan barang dan jasa melalui sistem elektronik	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • UU Perdagangan 2014 • PP PSTE 2019
konten yang mengandung elemen pelanggaran terhadap keamanan informasi	Pidana	<ul style="list-style-type: none"> • UU ITE 2008 jo. 2016 • UU Pertahanan Negara 2002 • UU Intelijen Negara 2011 • UU Keterbukaan Informasi Publik 2008 • PP PSTE 2019

konten yang mengandung elemen pelanggaran terhadap perlindungan konsumen	Pidana	<ul style="list-style-type: none"> ● UU ITE 2008 jo. 2016 ● UU Perlindungan Konsumen 199 ● PP PSTE 2019
konten yang mengandung elemen pelanggaran di sektor kesehatan	Pidana, Administratif	<ul style="list-style-type: none"> ● UU tentang Kesehatan (e.g., UU Praktek Kedokteran 2004, UU Kesehatan 2009, UU Rumah Sakit 2009) ● UU Narkotika 2009 ● UU ITE 2008 jo. 2016 ● PP PSTE 2019
konten yang mengandung elemen pelanggaran terhadap pengawasan obat dan pangan	Pidana, Administratif	<ul style="list-style-type: none"> ● UU Kesehatan 2009 ● UU Narkotika 2009 ● UU Pangan 2012 ● PP PSTE 2019
konten yang mengganggu ketertiban masyarakat dan publik (e.g. informasi dan/atau fakta yang salah)	Pidana	<ul style="list-style-type: none"> ● UU ITE 2008 jo. 2016 ● KUHP 1946 ● PP PSTE 2019

Sumber: Dikompilasi oleh penulis, 2021

Tabel di atas menunjukkan bahwa peraturan Indonesia tidak membedakan secara eksplisit antara konten “berbahaya” dan “ilegal”. Sebagai konsekuensinya, seluruh konten yang terdaftar di atas dapat diperlakukan sebagai konten “ilegal”, yang merupakan pelanggaran pidana secara substansi. Karena itu, setiap orang yang mempublikasikan, menyebarluaskan, dan mendistribusikan konten (yang memiliki potensi) berbahaya seperti yang tertulis di atas dapat dihukum secara pidana. Hal ini

merupakan konsekuensi logis dari absennya pembedaan eksplisit antara konten ilegal dan berbahaya di Indonesia.

Seperti yang telah dibahas, Laporan terhadap Pelapor Khusus tentang promosi dan perlindungan terhadap hak atas kebebasan berpendapat dan berekspresi (A/66/290) memberikan arahan untuk membedakan konten ilegal dan berbahaya. Di bawah Pelapor Khusus tersebut, konten ilegal berarti konten yang mengandung pelanggaran atas dasar hukum internasional dan dapat diadili secara pidana. Sementara itu, konten berbahaya dapat diklasifikasikan sebagai ekspresi yang tidak dapat dihukum secara pidana namun dapat menjadi dasar pembatasan atau gugatan sipil, dan ekspresi yang tidak mengakibatkan sanksi kriminal maupun sipil namun tetap menimbulkan kekhawatiran mengenai toleransi, kebermasyarakatatan, dan hormat terhadap satu sama lain.

Kekhawatiran mengenai pembedaan konten ilegal dan berbahaya juga dialamatkan di 1996 oleh EU melalui "Communication on illegal and harmful content on the Internet" (the Communication). Dokumen ini menekankan pentingnya membedakan antara konten daring ilegal dan berbahaya. Selain itu, baik Pelapor Khusus dan the Communication menekankan bahwa kategori konten yang berbeda memiliki "isu prinsip yang berbeda secara radikal, dan membutuhkan respon legal dan teknologis yang sangat berbeda".

Oleh karena itu, absennya pembedaan yang jelas antara konten ilegal dan berbahaya di peraturan Indonesia berpotensi menyebabkan permasalahan signifikan dalam pelaksanaan perlindungan hak-hak rakyat, terutama kebebasan berpendapat. Lebih dari itu, seperti yang telah dijelaskan di atas, kemungkinan untuk mengkriminalisasi ujaran yang bukan merupakan pelanggaran atau yang tidak dapat dihukum atas dasar standar internasional semakin besar.

Potensi terhadap permasalahan legal juga terlihat di berbagai putusan pengadilan daerah atas dasar setiap peraturan yang telah dibahas, contohnya Pengadilan Negeri Kebumen di 2018 yang memperluas objek

pencemaran nama baik dengan memasukkan kehormatan sebuah entitas legal atau institusi negara, Pengadilan Negeri Jakarta Selatan di 2019 yang mengelompokkan pendukung Kandidat Presiden dan Wakil Presiden dan penyelenggara pemilihan umum sebagai sebuah “kelompok”, Pengadilan Negeri Jayapura di 2020 yang menginterpretasi secara luas kebangsaan Indonesia sebagai bagian dari “kelompok” menurut Pasal 28 ayat (2) UU ITE, dan juga Pengadilan Negeri Kendari yang mengelompokkan kritik dan komentar negatif terhadap institusi negara sebagai ujaran kebencian atas dasar Pasal 28 (2) UU ITE. Hal ini menunjukkan bahwa ketentuan mengenai konten berbahaya dapat diinterpretasikan secara berbeda-beda. Maka dari itu, seberapa luas atau sempitnya sebuah konten dapat dikelompokkan sebagai berbahaya bergantung pada bagaimana pengadilan menginterpretasi ketentuan yang berhubungan dengan konten berbahaya.

Metode Penanganan Terhadap Konten Ilegal dan Berbahaya yang Berlaku Berdasarkan Regulasi di Indonesia

Seperti yang telah disebutkan, peraturan di Indonesia tidak membedakan konten ilegal dan berbahaya secara eksplisit, dan metode penanganan berbagai macam konten di atas cenderung sama. Dalam UU ITE, segala bentuk tindakan yang dilarang dapat dikenakan sanksi pidana. Hal yang sama juga berlaku untuk Kitab Undang-undang Hukum Pidana, UU Pornografi, dan UU tentang Penghapusan Diskriminasi Ras dan Etnis. Oleh karenanya, publikasi atau penyebaran konten tersebut dapat dikenakan sanksi pidana dan dihakimi di persidangan pidana.

Akan tetapi, hukum yang berlaku hanya menyediakan mekanisme pengadilan non-penal dan penyelesaian non-pengadilan atas dasar terbatas: (1) gugatan atas kerugian yang disebabkan penggunaan data pribadi secara non-konsensual; (2) gugatan terhadap penyelenggara sistem elektronik akibat kerugian yang mereka sebabkan; (3) gugatan

warga terhadap penyelenggara sistem elektronik yang membahayakan atau merugikan warga.⁴⁵ Meskipun beberapa mekanisme pengadilan non-penal melalui gugatan sipil dan penyelesaian non-pengadilan telah dijabarkan dalam UU ITE, penggunaan sarana ini tidak dapat mencabut tindakan kriminal yang telah dilakukan meskipun kasus tersebut sedang atau telah diadili.

Selain memberikan sanksi pidana terhadap pelanggar, Pasal 40 ayat (2a) UU ITE juga menjabarkan kewajiban pemerintah secara spesifik dalam mencegah konten terlarang berdasarkan hukum dan regulasi.⁴⁶ Lebih dari itu, menurut Pasal 40 ayat (2b), pemerintah diperbolehkan untuk memberhentikan akses dan/atau menginstruksikan Penyelenggara Sistem Elektronik untuk memberhentikan akses terhadap konten yang melanggar hukum dan regulasi,⁴⁷ dalam hal ini yaitu konten berbahaya yang dijabarkan di bagian sebelumnya. Di beberapa kasus, pengadilan dapat memerintahkan kementerian untuk memberhentikan akses terhadap konten tersebut.

Penyelenggara sistem elektronik (termasuk penyedia platform media sosial) diwajibkan memberhentikan akses terhadap konten tersebut berdasarkan instruksi Kementerian Komunikasi dan Informatika. Menurut Pasal 100 PP PSTE, mereka dapat dikenakan denda administratif dan pemerintah dapat memberhentikan akses terhadap konten ilegal dan berbahaya secara sepihak, bahkan memberhentikan akses ke platform tersebut. Mekanisme sepihak ini seringkali problematis karena ada kemungkinan perbedaan interpretasi mengenai konten apa yang dianggap ilegal dan berbahaya antara pemerintah dan platform. Ketentuan lanjutan terhadap permintaan pemberhentian dari publik, kementerian/agensi lain, kepolisian, jaksa, dan pihak yudisial, serta mekanisme pemberhentian akses terhadap konten terlarang yang diselenggarakan oleh Kementerian

⁴⁵ See Article 26, paragraph (2), Article 38, and Article 39 of Law No. 11 of 2008 jo. Law No. 19 of 2016 on Electronic Information and Transaction.

⁴⁶ Lihat Pasal 40 ayat (2a) UU No. 19 Tahun 2016 Tentang Amandemen UU No. 11 Tahun 2009 Tentang Informasi dan Transaksi Elektronik.

⁴⁷ Lihat Pasal 40 ayat (2b) UU No. 19 Tahun 2016 Tentang Amandemen UU No. 11 Tahun 2009 Tentang Informasi dan Transaksi Elektronik.

Komunikasi dan Informatika, juga oleh penyelenggara sistem elektronik dicantumkan dalam Peraturan Menteri Komunikasi dan Informatika tentang Penyelenggara Sistem Elektronik Lingkup Privat.

Selain mekanisme yang telah dijabarkan, beberapa undang-undang juga mengatur sarana spesifik berdasarkan peraturan sektor. Contohnya, dalam UU Pers, mekanisme penanganan dapat dilakukan melalui sensor dan pelarangan atau pembatasan. Sensor adalah penghapusan paksa terhadap sebagian atau seluruh materi informasi yang hendak dipublikasikan atau disiarkan, atau peringatan atau pemberitahuan bersifat intimidasi oleh pihak apapun, dan/atau kewajiban untuk melaporkan, dan memperoleh izin dari badan yang berwenang untuk melakukan kegiatan jurnalistik.⁴⁸

Sedangkan, pelarangan atau pembatasan terhadap penyiaran didefinisikan sebagai pemberhentian publikasi dan sirkulasi atau penayangan paksa yang melanggar hukum.⁴⁹ Kewenangan terhadap dua hal tersebut masih tidak jelas dalam UU Pers. Namun, secara prakteknya, penegakkan terhadap konten berbahaya, termasuk konten berbahaya yang diatur dalam UU Pers, dilakukan oleh Kementerian Komunikasi dan Informatika. Selain itu, perlu juga untuk merujuk kepada peraturan lain yang juga mengatur penempatan iklan dan peraturan lain yang berhubungan dengan obyek/materi/produk/jasa yang sedang diiklankan. Contohnya, Kementerian Kesehatan dapat meminta Kementerian Komunikasi dan Informasi untuk memblokir iklan rokok di media daring (Kementerian Kesehatan, 2019). Maka, menjadi penting untuk menetapkan mekanisme yang jelas dan terkoordinasi dalam penanganan konten pers di media online agar tidak terjadi tumpang tindih fungsi di lingkungan instansi terkait.

Selanjutnya, dalam UU Keterbukaan Informasi Publik dicantumkan ketentuan untuk pelanggaran terhadap konten yang diadakan atau

⁴⁸ Lihat Pasal 1 nomor 8 UU No. 40 Tahun 1999 tentang Pers.

⁴⁹ Lihat Pasal 1 nomor 9 UU No. 40 Tahun 1999 tentang Pers.

dipublikasi oleh badan umum. Pasal 19 UU Keterbukaan Informasi Publik menyatakan, "Pejabat Pengelola Informasi dan Dokumentasi di setiap Badan Publik wajib melakukan pengujian tentang konsekuensi sebagaimana dimaksud dalam Pasal 17 dengan saksama dan penuh ketelitian sebelum menyatakan Informasi Publik tertentu dikecualikan untuk diakses oleh setiap Orang."

Undang-undang lain yang menjelaskan mekanisme definitif secara spesifik adalah UU Hak Cipta. Dalam memproses konten ilegal dan berbahaya, UU Hak Cipta mencantumkan ketentuan untuk melaporkan pelanggaran hak cipta dalam sistem elektronik kepada Kementerian Hukum dan Hak Asasi Manusia, dan Menteri akan memverifikasi laporan tersebut. Selanjutnya, apabila terdapat bukti pelanggaran hak cipta yang cukup, Menteri akan merekomendasikan pemblokiran sebagian atau seluruh konten yang melanggar hak cipta dalam sistem elektronik atau memberhentikan akses terhadap jasa sistem elektronik tersebut kepada Kementerian Komunikasi dan Informatika.

Sebagaimana yang telah dijelaskan Pasal 56 (1) UU Hak Cipta, istilah 'memblokir konten dan/atau hak akses pengguna' berarti dua hal: Pertama, memblokir konten atau situs yang menyediakan jasa konten; Kedua, dalam bentuk memblokir akses pengguna terhadap situs spesifik dengan memblokir alamat protokol internet atau sejenisnya. Namun demikian, perlu dicatat bahwa tidak ada penjelasan atau alat ukur yang detail terhadap konten apa yang akan diblokir sebagian atau seluruhnya dan konten apa yang dapat memberhentikan akses ke jasa sistem elektronik. Selain itu, tidak dapat ditemukan ketentuan lebih lanjut mengenai sistem elektronik yang digunakan sebagai platform untuk membagikan media.

Secara keseluruhan, beberapa peraturan kunci yang mengatur platform media sosial di Indonesia dapat membahayakan kebebasan berpendapat secara daring melalui beberapa hal, seperti: memaksa platform media sosial untuk menghapus 'konten berbahaya' dengan memberikan jangka waktu yang terbatas dan ancaman sanksi; membatasi

metode moderasi konten dengan menjadikan penghapusan atau penurunan konten sebagai satu-satunya cara; mengurangi pilihan untuk memastikan praktek moderasi konten yang dilakukan secara transparan baik oleh pemerintah maupun platform media sosial.

Tanggung Jawab Platform Media Sosial dalam Mengatur Konten Ilegal dan Berbahaya

Beberapa peraturan mengatur kewajiban platform media sosial—atau didefinisikan secara legal sebagai Penyelenggara Sistem Elektronik (PSE)—untuk memoderasi konten di Indonesia. UU ITE dan peraturan pelaksanaannya—PP PSTE dan Peraturan Menteri Komunikasi dan Informatika No.5 Tahun 2020 (PP Menkominfo 5/20)—menjabarkan kewajiban platform untuk memoderasi konten dan konsekuensi apabila platform gagal mematuhi kewajiban tersebut. UU ITE memberikan gambaran umum tentang bagaimana platform harus bersikap dalam mengatur konten, dan hal ini semakin detail—bahkan semakin ketat dalam peraturan pelaksanaannya. Namun, SAFENet (2020) melaporkan bahwa dalam prakteknya, UU ITE umumnya digunakan terhadap individu (utamanya masyarakat biasa atau bagian dari masyarakat sipil) oleh mereka yang berada di posisi kekuasaan (pejabat pemerintahan dan bisnis).

Dalam PP PSTE, PSE atau platform diharapkan untuk menghapus konten ilegal dan konten yang dapat ‘menggangu ketertiban komunitas dan masyarakat’. Apabila gagal melakukan hal tersebut, PSE dapat dikenakan hukuman sesuai peraturan yang berlaku. Kewajiban PSE secara lebih detail dijelaskan di PP Menkominfo 5/20. Peraturan ini mengandung ketentuan mengenai batas waktu yang diberikan kepada PSE untuk merespon permintaan pemerintah untuk menghapus konten (24 jam setelah pemberitahuan atau 4 jam dalam kasus permintaan penghapusan konten yang ‘mendesak’, tanpa penjelasan mengenai situasi apa yang dapat disebut ‘mendesak’). Selain itu, peraturan ini memberikan sanksi terhadap PSE yang gagal melakukan tindakan tersebut dalam batas waktu yang diberikan. Bentuk sanksi beragam, mulai dari pemblokiran akses hingga denda.

PP Menkominfo 5/20, yang merupakan peraturan tentang moderasi konten yang relatif lebih detail, problematis karena beberapa hal. Contohnya, peraturan ini tidak memberikan proses hukum untuk permintaan penghapusan, terutama permintaan dari pemerintah (Pasal 19, 2021). Batas waktu yang pendek tidak memberikan waktu bagi PSE untuk menilai permintaan penghapusan konten secara seksama. Pada akhirnya, PSE dapat dipaksa untuk tunduk demi menghindari hukuman administratif-hukum ini bersifat punitif secara prinsip.

Peraturan Menteri ini diduga memberikan kekuasaan lebih kepada pemerintah untuk mengatur informasi yang disebarluaskan dan memberlakukan sensor di internet (Pasal 19, 2021). Peraturan ini menekankan penghapusan atau penghilangan konten sebagai satu-satunya pilihan yang mungkin untuk moderasi konten. Diikuti dengan batas waktu yang pendek, peraturan ini dapat digunakan untuk menyensor ungkapan yang sah dan dapat membahayakan demokrasi dalam jangka panjang. Dalam aspek yang lebih praktis, penghapusan sebagai satu-satunya pilihan untuk moderasi konten tidak sejalan dengan mekanisme moderasi konten yang dimiliki kebanyakan PSE, terutama yang digunakan untuk mengatur ungkapan di platform mereka.

Meskipun PP Menkominfo ini cenderung masih baru, Google melaporkan bahwa Indonesia berada di urutan 10 peringkat teratas dalam permintaan penghapusan konten dan memimpin di peringkat pertama dalam hal penghapusan konten (CNN Indonesia, 2021a). David Graff, Wakil Presiden bidang Keamanan dan Kepercayaan Google, menyebutkan bahwa mereka berusaha untuk mematuhi perintah pengadilan yang berlaku di masing-masing negara. Dalam kasus Indonesia, hampir semua permintaan yang berasal dari pemerintah berlandaskan alasan keamanan nasional (CNN Indonesia, 2021a).

Pada 2019, juru bicara Kementerian Komunikasi dan Informatika menyatakan bahwa Facebook sangat keras menanggapi permintaan penghapusan dari pemerintah, menyebutkan bahwa mereka selalu tidak

mengindahkan permintaan tersebut, mengutip interpretasi yang berbeda mengenai apa yang dapat disebut konten terlarang (Farras, 2019). Dalam white paper yang diterbitkan oleh Facebook (yang dijelaskan secara lebih detail di bawah), mereka menawarkan kerjasama sembari memberi saran untuk kebijakan pemerintah yang lebih spesifik dan mengkritisi paradigma yang memberatkan penghapusan konten dan kuantitas dalam moderasi konten (Bickert, 2020).

Penghapusan konten cukup umum selama hal ini diminta oleh pemerintah. Sementara, ajuan untuk penghapusan konten dari pengguna pada umumnya jarang didengar. Dalam riset kolaboratif antara The University of Sydney dan The University of Queensland (Sinpeng et al. 2021)-yang didanai oleh Facebook-, ditemukan bahwa individu cenderung mengurungkan niat untuk melaporkan apa yang mereka anggap konten terlarang, terutama ujaran kebencian dan konten negatif lain terhadap minoritas dan kelompok rentan, karena persepsi mereka terhadap minimnya dampak dari praktek moderasi Facebook.

Namun, meskipun peraturan yang ada memaksa platform untuk mematuhi permintaan penghapusan, ketentuan yang memandatkan transparansi dari proses keseluruhan kepada badan pemerintahan maupun platform masih kurang. Platform media sosial korporat seringkali mengabstraksi praktek dan kebijakan tekno-sosial mereka, sehingga hanya sedikit hingga tidak ada sama sekali informasi mengenai moderasi konten yang sampai kepada orang awam. Peraturan yang ada di Indonesia juga kurang dalam mengatur mandat terhadap platform atau badan pemerintahan untuk membuka seluruh proses dari praktek moderasi konten sesuai permintaan pemerintah. Saat ini, platform media sosial telah mulai mempublikasikan laporan transparansi mandiri dalam berbagai bentuk, namun seringkali laporan ini hanya berfokus pada jumlah 'konten berbahaya' yang telah sukses dimoderasi selama setahun dibandingkan bagaimana cara konten tersebut dimoderasi. Pendekatan yang sama diambil oleh pemerintah—meskipun lebih mengkhawatirkan karena penghapusan postingan dan situs dianggap sebagai capaian. Dalam

memastikan moderasi konten yang tidak mengkompromisi kebebasan berpendapat, keterbukaan angka tidak cukup dan laporan transparansi yang lebih detail dan mendalam menjadi sangat vital.

Sinpeng dkk. (2001) merekomendasikan Facebook (dan perusahaan media sosial lainnya) untuk lebih transparan dalam praktek moderasinya (dari atas ke bawah atau top to bottom), memberdayakan administrator pengguna dan tidak hanya moderator konten internal atau outsourcing, serta melibatkan auditor pihak ketiga.

Selain itu, meskipun Facebook dan perusahaan media sosial lain menawarkan dokumen publik tentang kebijakan moderasi masing-masing, mereka memiliki dokumen internal yang mengarahkan moderasi konten. Bentuk dokumen ini yang lebih sederhana dan user-friendly dibutuhkan demi pelaksanaan moderasi konten yang bertanggung jawab. Lebih dari itu, dokumen ini harus dipublikasikan dalam berbagai bahasa daerah sebanyak mungkin. Hal ini dapat dilakukan dengan berkoordinasi dan bekerjasama dengan penduduk lokal atau bahkan anggota komunitas rentan yang terorganisir. Pengguna harus memiliki akses terhadap kebijakan penalti dan proses banding yang tidak atau sulit untuk diketahui.

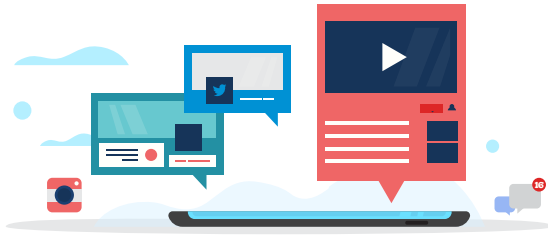




BAB II

Tren





Pembahasan mengenai tren pelaksanaan peraturan konten berbahaya di Indonesia dapat dianalisis melalui berbagai pendekatan. Digunakan dua perspektif utama di riset ini: pertama, melalui lensa pemerintah dan negara sebagai regulator; dan kedua, melalui perspektif warga sebagai kelompok yang terdampak. Perspektif pertama mendalami ketentuan yang diterbitkan oleh institusi pemerintah dalam merespon isu-isu terkait konten di media daring, sementara perspektif kedua menyorot berbagai kasus yang terjadi di masyarakat. Selain itu, isu-isu terkait teknologi dan moderasi konten otomatis juga akan didiskusikan untuk memahami potensi implikasi dari penggunaan teknologi dalam menyuarakan perdamaian dan meningkatkan ketahanan masyarakat terhadap konten berbahaya.

Kebijakan yang Dikeluarkan dalam Menanggapi Masalah Terkait Konten di Masyarakat

Meskipun banyak hukum dan peraturan telah dilaksanakan, masih banyak permasalahan yang tersisa akibat pelaksanaan UU ITE dan peraturan lainnya yang berhubungan yang menimbulkan interpretasi yang beragam oleh polisi, jaksa, dan pengadilan dan kontroversi di masyarakat (Lihat, contohnya, Evandio, 2021; Rahmawati, 2021). Kekhawatiran institusi pemerintah mengenai pelaksanaan ketentuan UU ITE dan peraturan sehubungan yang multitafsir dapat dilihat, contohnya, pada Surat Keputusan Bersama Pedoman Kriteria Implementasi UU Informasi dan Elektronik (UU ITE) yang ditandatangani oleh Menteri Komunikasi dan Informatika, Kepala Kepolisian Negara Republik Indonesia (Kapolri), dan

Jaksa Agung. Oleh karenanya, di beberapa tahun terakhir, institusi-institusi negara dan pemerintah telah mengeluarkan banyak kebijakan mengenai pelaksanaan peraturan yang berhubungan dengan konten sebagai respon terhadap kondisi masyarakat. Meskipun begitu, hampir semua kebijakan yang dikeluarkan pada intinya ditujukan sebagai pedoman dalam mengimplementasikan berbagai peraturan yang telah dijelaskan di bab sebelumnya. Akan tetapi, tidak dapat dipungkiri bahwa beberapa ketentuan di peraturan-peraturan ini cenderung multitafsir dan membahayakan kebebasan berpendapat. Penjelasan mengenai beberapa peraturan yang relevan akan dijabarkan di bawah.

Surat Edaran Kepala Kepolisian Negara

→ Republik Indonesia Nomor SE/6/X/2015 tentang Penanganan Ujaran Kebencian

Surat edaran ini ditujukan sebagai pedoman bagi polisi dalam menghadapi kasus ujaran kebencian dan maraknya kekerasan, terutama konflik agama yang terjadi di 2015. Contohnya, pembakaran masjid di Tolikara (Jakarta Globe, 2015) dan pembakaran gereja di Aceh Singkil (Dewi Suci Rahayu, 2015) sebagai respon terhadap pesan-pesan provokatif dan spekulatif yang beredar secara daring, tidak terkecuali pemilihan umum Indonesia pada Desember 2015 (Zaqiu Rahman, 2015). Surat edaran ini penting, terutama bagi anggota kelompok agama minoritas yang haknya telah dilanggar di ruang digital.

Surat edaran ini mendefinisikan ujaran kebencian dengan menyebut beberapa ketentuan yang sudah ada di dalam hukum pidana sembari memperluas cakupan pelanggarannya. Saat ini, ujaran kebencian termasuk penghinaan; pencemaran nama baik; penistaan; perbuatan tidak menyenangkan; memprovokasi; menghasut, dan penyebaran berita bohong; dimana semua tindakan tersebut ditujukan atau berakibat pada tindakan diskriminasi; kekerasan; kehilangan nyawa; dan/atau konflik sosial.

Surat edaran ini juga membedakan target ujaran kebencian berdasarkan beberapa aspek komunitas seperti: suku; agama; aliran keagamaan; keyakinan/kepercayaan; ras; antargolongan; warna kulit; etnisitas; gender; kaum difabel (cacat); orientasi seksual.

Lebih dari itu, surat edaran ini juga memberikan pendekatan preventif dan represif dalam menghadapi kasus ujaran kebencian. Pendekatan preventif menekankan ajakan kepada anggota kepolisian untuk memahami tiap bentuk ujaran kebencian, sehingga meningkatkan penanganan kejahatan yang lebih efisien dan berhati-hati. Sebaliknya, pendekatan represif berfokus pada cara menegakkan hukum.

Beberapa aktivis, kelompok media, dan pengacara merespon surat edaran ini dengan mengkritisi bahwa surat edaran ini dapat berdampak pada pelanggaran hak asasi manusia, yaitu kebebasan berpendapat (Mong Palatino, 2015). Sebagaimana yang telah disebutkan, surat edaran ini mengkategorikan penghinaan, pencemaran nama baik, dan penistaan sebagai bagian dari ujaran kebencian. Meskipun demikian, hal ini dapat berakibat pada salah tangkap karena ia tidak menjabarkan secara jelas definisi dan karakteristik tindakan-tindakan tersebut. Sebaliknya, surat ini hanya merujuk pada ketentuan di peraturan-peraturan yang sudah ada yang sebelumnya telah disebutkan di bab terpisah, sehingga dapat menciptakan interpretasi yang beragam. Tujuan utama dari penerbitan surat edaran ini sebenarnya patut diapresiasi karena pemerintah telah mulai mengakui adanya ujaran kebencian, terutama di ruang digital. Namun begitu, tidak dapat dipungkiri bahwa beberapa pernyataan di surat edaran ini dapat menyebabkan dan melegitimasi pengambilan keputusan dan penyimpangan penegakkan hukum yang subyektif, terutama terhadap kebebasan berpendapat dan berekspresi.

Surat Telegram Kepala Kepolisian Negara Republik Indonesia Nomor ST/1100/IV/HUK.7.1./2020 4/04/2020

→ tentang Bentuk Pelanggaran atau Kejahatan serta Masalah yang Mungkin Terjadi dalam Perkembangan Situasi serta Opini di Ruang Siber

Latar belakang utama dibalik penerbitan surat telegram ini adalah respon terhadap pembatasan sosial berskala besar (PSBB) di tengah implementasi kebijakan COVID-19 (Adi Briantika, 2020). Sejak Januari 2020, Kementerian Komunikasi dan Informatika telah mencatat 1.971 hoax yang berhubungan dengan COVID-19 pada 5.065 unggahan di media sosial (WM, 2021). Surat telegram ini diharapkan dapat membantu pemerintah dalam menegakkan implementasi kebijakan PSBB karena penyebaran misinformasi dan disinformasi terkait COVID-19 di ruang digital. Surat telegram ini memandu polisi untuk melaksanakan patroli siber yang terpadu untuk mengawasi situasi dan opini mengenai disinformasi tentang COVID-19, langkah pemerintah dalam menangani COVID-19, dan penghinaan terhadap Presiden dan/atau pejabat-pejabat lainnya. Secara umum, surat ini menghasilkan ketentuan-ketentuan operasional bagi polisi dalam menangani kasus-kasus tersebut.

Surat telegram ini memerintahkan penyelidik kepolisian untuk mengambil tindakan hukum atas setiap pelanggar yang berhubungan dengan disinformasi mengenai COVID-19. Namun demikian, beberapa melayangkan kritik sebagai respon terhadap penerbitan surat telegram ini, termasuk mempertanyakan apabila surat telegram ini memiliki potensi untuk membatasi—atau bahkan melanggar kebebasan berpendapat, beropini, dan kebebasan akademik. Lokataru Foundation—gerakan dan advokat hak asasi manusia (HAM) dan dasar hukum di Indonesia—menyebutkan bahwa selain mengatur ‘patroli siber’ polisi untuk melawan ‘berita palsu’ COVID-19, surat telegram ini juga meregulasi penghinaan terhadap Presiden dan/atau pejabat pemerintahan lainnya (Lokataru Foundation, 2020). Hal ini

menyebabkan stigma publik yang menganggap bahwa surat telegram ini dapat digunakan untuk membungkam dan membatasi kritik terhadap pemerintah alih-alih memonitor hoax tentang COVID-19 (BEM KEMA Universitas Padjajaran, 2020). Implementasi surat telegram ini meningkatkan ancaman dan penangkapan terhadap aktivis hak asasi manusia (HAM) karena mengkritik respon dan kebijakan negara terhadap COVID-19. Lebih dari itu, Amnesty International juga mengajukan untuk menarik kembali surat telegram ini secepatnya karena potensinya untuk memberikan ruang bagi pihak berwajib untuk menyalahgunakan kekuasaannya terhadap kebebasan berpendapat (BEM KEMA Universitas Padjajaran, 2020).

Lebih dari itu, surat telegram ini memperbolehkan penyelidikan kepolisian untuk menggunakan alasan 'keadaan darurat COVID-19' sebagai alasan untuk menjustifikasi tindakan mereka dalam proses penegakan hukum karena tidak adanya peraturan atau kebijakan yang secara eksplisit dan spesifik menyatakan ukuran dan kategorisasi penghinaan: apakah yang dapat dikatakan sebagai penghinaan?

Sebagaimana yang telah dijabarkan, respon terhadap penerbitan surat telegram ini menganggap bahwa surat telegram ini dapat mengakibatkan pelanggaran hak pers dan kebebasan berpendapat masyarakat, terutama mengenai COVID-19. KontraS, sebuah komisi yang bergerak untuk mengawasi isu hak asasi manusia, menganggap penerbitan surat telegram ini sebagai hal yang berpotensi mengancam hak masyarakat untuk menyuarakan opini mereka (terutama mengenai isu tentang COVID-19) dan mengakui potensinya dalam mendorong penyalahgunaan kekuasaan yang lebih luas oleh aparat kepolisian (Harian Jogja, 2022).



Surat Edaran Kepala Kepolisian Negara Republik Indonesia Nomor 2/11/2021 tentang Kesadaran Budaya Beretika untuk Mewujudkan Ruang Digital Indonesia yang Bersih, Sehat, dan Produktif

Surat edaran ini ditujukan untuk memperjelas UU No.19/2016 tentang amandemen UU No. 2008 tentang ITE, mempertimbangkan bagaimana peraturan yang ada tidak menjelaskan secara langsung maupun spesifik mengenai cara pelaksanaannya. Selain itu, surat ini juga diedarkan untuk menjelaskan kekhawatiran dalam pengaplikasian UU ITE, yang berhubungan dengan potensi untuk mengkriminalisasi dan melaporkan beberapa pihak. Surat edaran ini 'diharapkan' dapat menjamin hak-hak komunitas digital, termasuk hak kebebasan berpendapat. Secara umum, surat edaran ini dapat digunakan untuk mengarahkan polisi dan jaksa dalam menangani pelanggaran terhadap UU ITE, termasuk konten ilegal dan berbahaya di ruang digital.

Setelah penerbitan surat edaran ini, Kepolisian Negara Republik Indonesia merilis polisi virtual yang berfokus pada pencegahan melalui pengawasan, edukasi, peringatan, dan pencegahan individu untuk tidak melakukan tindakan kriminal di ruang digital. Penerbitan surat edaran ini mendapat respon yang sama dengan kebijakan sebelumnya, yang menganggap bahwa kebijakan ini dapat mengancam kebebasan berpendapat. Seperti yang dinyatakan oleh SAFEnet, hal ini dapat diinterpretasikan sebagai usaha kepolisian untuk menciptakan 'panoptisisme digital'.

Keberadaan polisi virtual dapat berpotensi menyebabkan opresi teknologis karena kekuasaannya untuk melakukan sensor secara daring, pengawasan siber, dan usaha untuk mengontrol sebuah infrastruktur (Nugraha dan Laila, 2021). Menurut KontraS, surat edaran ini hanya mengatur pembentukan polisi virtual, sementara mekanisme kontrol dan keterbukaan informasi publik masih belum diatur. Oleh

karena itu, ketiadaan transparansi dari operasi polisi virtual mengakibatkan surat edaran ini menjadi multitafsir bagi pihak berwajib. Di dua bulan pertama operasi polisi virtual ini, terdapat 329 konten yang dianggap sesuai dengan UU ITE dan 200 di antaranya lolos proses verifikasi. Menurut KontraS, polisi menurunkan konten yang juga mengandung kritik terhadap pemerintah saat operasi polisi virtual ini yang memicu kemarahan publik. (KontraS, 2021). Beberapa laporan di atas menunjukkan bahwa penerbitan surat edaran ini sebagai dasar hukum pendirian polisi virtual dapat juga berpotensi mengakibatkan pelanggaran terhadap hak kebebasan berpendapat (KontraS, 2021).

➔ **Surat Telegram Kapolri No. ST/339/II/RES.1.1.1./2021 tentang Pedoman Penanganan Perkara Tindak Pidana Kejahatan Siber**

Surat edaran ini diterbitkan sebagai pedoman bagi seluruh penyelidik kepolisian dalam menangani kasus kejahatan siber, terutama ujaran kebencian. Melalui peraturan-peraturan ini, keadilan restoratif dapat digunakan dalam proses pengadilan pidana yang berhubungan dengan pelanggaran terhadap Pasal 27 (3) UU ITE tentang penghinaan dan/atau pencemaran nama baik; Pasal 270 (fitnah), 310 (pencemaran nama baik), dan 311 (pencemaran nama baik) Kitab Undang-undang Hukum Pidana (KUHP) Indonesia. Pada dasarnya, surat telegram ini ditujukan untuk membantu penyelidik kepolisian dalam memaksimalkan penegakkan hukum terhadap pelaku kejahatan siber.

Pelaksanaan surat telegram ini dapat ditemukan pada kasus Novel Baswedan dan kasus Roy Suryo. Novel Baswedan dilaporkan oleh Pemuda, Pelajar, dan Mahasiswa Mitra Kamtibmas (PPMK) atas tweet-nya mengenai kematian Soni Erata alias Ustadz Maaher At-Thuwailibi (VOI, 2021). Kepolisian Negara Republik Indonesia menyebutkan bahwa sehubungan dengan yang tertera pada surat telegram ini, mediasi sebagai bentuk keadilan restoratif digunakan untuk

menangani kasus Novel Baswedan (Antara dan Kukuh, 2021). Sama dengan kasus Novel Baswedan, Roy Suryo sebelumnya melaporkan Lucky Alamsyah dengan tuduhan pencemaran nama baik. Lucky Alamsyah dituduh menyebarkan hoax dalam unggahannya yang membahas masalah kecelakaan lalu lintas yang ia alami dengan Roy Suryo (VOI, 2021). Kasus ini ditutup secara damai melalui mediasi.

Penerbitan surat edaran ini perlu diapresiasi karena surat ini memberikan ruang untuk memprioritaskan keadilan restoratif dan menangani beberapa kejahatan di ranah siber. Surat telegram ini mengajak pihak berwajib untuk mengimplementasikan keadilan restoratif sembari menangani kasus ujaran kebencian. Hal ini juga dapat digunakan untuk menentukan kejahatan siber yang dapat ditangani melalui keadilan restoratif karena surat telegram ini menyatakan dengan eksplisit bahwa penting untuk memberikan kepastian hukum. Sebagaimana yang telah tercermin di kasus-kasus yang dibahas di atas, keduanya memprioritaskan penggunaan dua pendekatan yang berbeda untuk menangani konflik yang terjadi di ranah siber—melalui mediasi—yang memberikan pihak-pihak terlibat kesempatan untuk bertemu dan berkomunikasi untuk memperbaiki hubungan dan mengurangi proses yudisial tradisional yang menghabiskan uang dan waktu.

Surat Keputusan Bersama Menteri Komunikasi dan Informatika, Jaksa Agung, Dan Kapolri Nomor 229

→ Tahun 2021; 154 Tahun 2021; KB/2/VI/2021 Tahun 2021 tentang Pedoman Implementasi atas Pasal Tertentu dalam UU ITE

Surat Keputusan Bersama (SKB) ini ditujukan untuk menjaga ruang digital Indonesia yang bersih, sehat, etis, dan produktif. SKB ini akan memandu Kementerian Komunikasi dan Informatika, Kejaksaan Agung, dan Kepolisian Negara Republik Indonesia dalam menjalankan tugas dan wewenangnya untuk melaksanakan penegakkan hukum

yang berhubungan dengan pelanggaran terhadap UU ITE. SKB ini berusaha menyatukan interpretasi dari beberapa ketentuan sehubungan dengan konten yang dilarang oleh UU ITE.

Latar belakang utama SKB ini adalah kekhawatiran masyarakat mengenai pasal-pasal karet dalam UU ITE yang sering digunakan untuk kriminalisasi (Rosana dan Eko, 2021). Belum lagi, keberadaan beberapa ketentuan yang mengakibatkan beberapa interpretasi dapat membingungkan polisi, jaksa, dan pengadilan. Tujuan utama dari SKB ini patut diapresiasi. Namun demikian, SKB ini mendapat kritik karena proses penyusunannya yang tidak terbuka dan tidak melibatkan partisipasi masyarakat (Budiarti Utami Putri, 2021). Lebih dari itu, SKB ini hanya sekadar pedoman internal bagi Kementerian Komunikasi dan Informatika, Jaksa Agung, dan Kepolisian Negara Republik Indonesia. Koalisi Serious Revisi UU ITE menyebutkan bahwa hal yang paling krusial untuk dilakukan terlebih dulu adalah amandemen seluruh pasal-pasal karet dalam UU ITE.

Tidak dapat dipungkiri bahwa proses pengadilan pidana dimulai di tingkat kepolisian, yang merupakan pintu masuk ke proses yudisial. Pada intinya, seluruh kebijakan yang telah disebutkan di atas, yaitu surat edaran; surat telegram; dan surat keputusan bersama (SKB) diterbitkan untuk memandu institusi bersangkutan, terutama kepolisian, dalam menangani proses penegakan hukum pada kasus konten berbahaya, seperti ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik. Akan tetapi, sebagaimana yang telah dinyatakan di awal sub-bab ini, terdapat ruang untuk perbaikan terhadap beberapa peraturan, terutama peraturan multitafsir antara polisi, jaksa, dan pengadilan.

Tren Berdasarkan Perspektif Masyarakat

Ketentuan mengenai konten ilegal dan berbahaya, terutama dalam UU ITE, telah menjadi diskursus publik dan sudah lama dikritik oleh berbagai organisasi masyarakat sipil sejak UU ITE ditetapkan pada tahun 2008.⁵⁰ Berbagai pihak juga telah berulang kali mengajukan uji materi terhadap UU ITE ke Mahkamah Konstitusi.⁵¹ Dari seluruh klasifikasi yang telah ditulis di bab sebelumnya, terdapat tiga kategori yang paling sering dijadikan dasar gugatan dan dimanfaatkan oleh penegak hukum dalam penuntutan pidana, yaitu (1) fitnah atau pencemaran nama baik, (2) ujaran kebencian, dan (3) konten yang bertentangan dengan kesusilaan. Selain itu, penggunaan ketentuan mengenai misinformasi dan disinformasi telah meningkat secara bertahap dalam beberapa tahun terakhir (Lihat, contoh., Tempo, 2020; Debora, 2020; dan SAFEnet, 2021).

Selanjutnya, laporan SAFEnet mengenai hak digital di Indonesia mencatat adanya 84 kasus kriminalisasi terhadap pengguna internet di Indonesia pada tahun 2020, sebuah peningkatan pesat dibandingkan tahun sebelumnya dengan 24 kasus. UU ITE tetap menjadi peraturan utama untuk membatasi suara masyarakat. Ketentuan-ketentuan yang digunakan adalah ujaran kebencian (27 kasus), pencemaran nama baik (22 kasus), dan disinformasi dalam transaksi elektronik (12 kasus). Selain UU ITE, peraturan yang digunakan antara lain KUHP, khususnya Pasal 14 dan Pasal 15 yang mengatur tentang misinformasi dan disinformasi, serta Pasal 270, 310, dan 311 yang meliputi fitnah dan pencemaran nama baik. Diskursus publik pun tetap sama karena kedua peraturan tersebut, sebagaimana diimplementasikan, tetap bergantung pada kondisi kebebasan berpendapat dan berekspresi yang ada.

Menariknya, menurut laporan-laporan dari Tirto dan Detik, pihak yang paling sering menggunakan pasal-pasal tersebut adalah pejabat publik

⁵⁰ Southeast Asia Freedom of Expression Network (SAFEnet) telah menerbitkan laporan dan rilis terkait kasus UU ITE. Lihat laporan dan berita mengenai UU ITE di <https://id.safenet.or.id/?s=UU+ITE>, dan daftar kasus UU ITE di <https://id.safenet.or.id/daftar kasus/daftar kasus tidak diperbarui sejak Desember 2020>.

⁵¹ Lihat misalnya, Mahkamah Konstitusi Republik Indonesia, Putusan Mahkamah Konstitusi, <https://www.mkri.id/index.php?page=web.Putusan&id=1&kat=2&cari=Informasi+dan+Transaksi+Elektronik>

(Lihat Detik, 2021; Tirto, 2018). Tirto (2018) mencatat bahwa 35,92% pelapor kasus UU ITE adalah pejabat publik, termasuk kepala daerah, pejabat kementerian, dan penegak hukum. Dalam laporannya, Tirto menambahkan bahwa pengajuan kasus UU ITE oleh pejabat publik atau pemerintah sebagian besar disebabkan oleh ekspresi dan kritik terhadap kinerja atau posisi para pejabat tersebut.

Hal tersebut pun ditegaskan dan meningkat pada tahun 2021, di mana berdasarkan laporan Detik, 70% kasus UU ITE yang dilaporkan dari tahun 2017 hingga 2019 diajukan oleh pihak-pihak yang memiliki kekuasaan, termasuk pejabat, pengusaha, bahkan polisi sendiri. Sedangkan 29% lainnya dilaporkan oleh warga. Selanjutnya, menurut laporan SAFEnet pada tahun 2021, dari 84 subjek yang dilaporkan, 50 merupakan warga sipil, 15 aktivis, empat buruh, tiga pegawai swasta, dua mahasiswa, dan seseorang jurnalis. Mereka menjadi korban tuduhan pencemaran nama baik, fitnah, penyebaran permusuhan antar kelompok dan ujaran kebencian, serta misinformasi/disinformasi karena penggunaan pasal-pasal tersebut untuk melindungi dan menutupi pejabat publik dan kebijakan-kebijakan mereka.

Sebuah koalisi organisasi masyarakat sipil di Indonesia pun menyusun sebuah laporan yang mendesak beberapa reformasi terhadap regulasi konten dan rezim moderasi konten di Indonesia (Amnesty International Indonesia et al., 2021). Koalisi tersebut berpendapat bahwa bentuk regulasi konten saat ini, khususnya UU ITE, dan implementasinya seringkali berpihak pada pelanggaran kebebasan berekspresi, yang sangat jauh dari menjamin perdamaian dan stabilitas bangsa. Undang-undang tersebut dianggap memberikan terlalu banyak kekuasaan dan keleluasaan kepada penegak hukum tanpa proses hukum yang semestinya dan langkah-langkah akuntabilitas yang kuat. Data koalisi tersebut menunjukkan bahwa 96.8% dari orang-orang yang dituntut dengan UU ITE dijatuhi hukum pidana, dan 88% dijatuhi hukuman pidana penjara; keduanya dianggap sangat tinggi.

Koalisi yang terdiri dari 24 organisasi yaitu Amnesty International Indonesia, Aliansi Jurnalis Independen, Greenpeace Indonesia, dan SAFEnet tersebut menemukan bahwa UU ITE terkadang digunakan oleh aktor-aktor berpengaruh, seperti pejabat pemerintah dan pengusaha, untuk melindungi kepentingan mereka. Salah satu kasus penting yang menggambarkan temuan tersebut adalah kasus Ravio Patra, seorang peneliti, yang dekriminalisasi menggunakan Pasal 27 ayat (3) UU ITE oleh Wempy Dyocta Koto, seorang motivator bisnis, atas dasar pencemaran nama baik karena sebuah unggahan di Facebook (Widhana, 2017). Wempy adalah seorang tokoh masyarakat yang sering memberikan nasihat bisnis di media sosial dan mengklaim dirinya memegang beberapa penghargaan. Dengan skeptis, Ravio pun meneliti latar belakang Wempy dan menemukan beberapa inkonsistensi. Sebagai contoh, Wempy mengatakan bahwa ia memegang Penghargaan Kewirausahaan Tertinggi Asia dan Penghargaan Kepemimpinan Tertinggi Asia (Asia's Highest Entrepreneurship Award and Asia's Highest Leadership Award), namun kedua penghargaan tersebut sebetulnya tidak ada. Wempy kemudian mengklarifikasi bahwa ia sebenarnya menerima Penghargaan Kewirausahaan Asia Pasifik (Asia Pacific Entrepreneurship Award) pada tahun 2013 dan Penghargaan Perusahaan Keunggulan dan Keberlanjutan Asia (Asia Corporate Excellence and Sustainability Award) pada tahun 2016. Baik Ravio dan Wempy pun berkorespondensi secara daring, dengan pertanyaan dan jawaban yang mengalir di antara keduanya. Namun pada 21 Juni 2017, Ravio dituduh melakukan pencemaran nama baik dan menyebabkan kerugian finansial bagi Wempy. Tidak hanya itu, Ravio juga mendapat tuntutan aneh dari Wempy, seperti ganti rugi Rp 5 miliar, menulis surat permintaan maaf untuk dipasang di media, media sosial, dan dalam bentuk video meskipun Ravio belum pernah mengunggah video apa pun di Youtube.

Selain itu, ketidakjelasan penjelasan mengenai subjek pencemaran nama baik dan cakupan konten pencemaran nama baik juga menimbulkan polemik dalam pelaksanaannya. Sebagai contoh, pada tahun 2010, Mahkamah Agung menyatakan bahwa kehormatan badan hukum,

termasuk lembaga negara, dapat menjadi objek konten fitnah.⁵² Dengan mengacu pada Putusan Mahkamah Agung tahun 2010 tersebut, pada tahun 2018, Syaeful Lillah dinyatakan bersalah oleh Pengadilan Negeri Kebumen karena melakukan pencemaran nama baik terhadap Kepolisian Negara (Polri) berdasarkan Pasal 27 ayat (3) UU ITE.⁵³ Putusan tersebut dapat menggambarkan bagaimana UU ITE dapat digunakan oleh pemerintah dan lembaga negara lainnya untuk mengkriminalisasi masyarakat.

Contoh kasus lain adalah ketika Sadli Saleh, pemimpin redaksi liputanpersada.com, menulis editorial yang mengkritik 'proyek janggal yang dilakukan oleh pemerintah Buton Tengah'. Sadli Saleh dikriminalisasi menggunakan Pasal 28 ayat (2) yang meliputi hasutan kebencian atau permusuhan (Koran Tempo, 2020). Sadli melaporkan, proyek simpang lima itu menggelembung di bagian anggaran, dari Rp4 miliar menjadi Rp6.8 miliar, tetapi lajur yang diproyeksikan lima berkurang menjadi empat. Terlebih lagi, Sadli juga mengkritik kurangnya transparansi dan kurang baiknya perencanaan proyek tersebut, yang ditunjukkan dengan alokasi anggaran proyek yang membengkak. Kesal dengan laporan tersebut, pejabat pemerintah melaporkan jurnalis tersebut kepada pihak berwenang atas dasar pencemaran nama baik dan hasutan untuk permusuhan. Padahal, karena laporan tersebut merupakan sebuah karya jurnalistik, pejabat tersebut seharusnya bisa menggunakan hak untuk menjawab (right to reply) (Bernie, 2020). Contoh lain juga dapat dilihat dalam Putusan PN Kendari No. 426/Pid.Sus./2021/PN Kdi, di mana MK berpendapat bahwa kritik keras dan komentar negatif terhadap lembaga negara dapat tercakup dalam Pasal 28 ayat (2) UU ITE jika orang tersebut memahami bahwa akan ada berbagai tanggapan terhadap konten tersebut, yang menyebabkan polemik sosial melalui media sosial.⁵⁴ Namun, tidak ada penjelasan yang jelas dari MK mengenai apa yang bisa dianggap sebagai polemik sosial.

⁵² Lihat Putusan Mahkamah Agung No. 1833K/Pid./2010 (20 Mei 2010) 15.

⁵³ Lihat Putusan Pengadilan Negeri Kebumen No. 223/Pid.Sus./2018/PN Kbm (17 Desember 2018).

⁵⁴ Lihat Putusan Pengadilan Negeri Kendari No. 426/Pid.Sus./2021/PN Kdi (16 September 2021) 24.

Pada tahun 2021, seorang tokoh masyarakat Indonesia, Jerinx, juga dikriminalisasi berdasarkan Pasal 28 ayat (2) UU ITE tentang penyebaran informasi yang bertujuan menimbulkan kebencian atau permusuhan kepada individu dan/atau kelompok masyarakat tertentu berdasarkan suku, agama, ras, dan antar kelompok.⁵⁵ Kasus ini berawal dari unggahan Jerinx di akun media sosial Instagram-nya yang mengatakan bahwa Ikatan Dokter Indonesia (IDI) adalah 'pelemah Organisasi Kesehatan Dunia (WHO)' Dalam putusannya, hakim menyatakan IDI termasuk dalam kelompok 'antargolongan' yang dilindungi Pasal 28 ayat (2) UU ITE. Namun, keputusan hakim tersebut dinilai kurang tepat oleh beberapa pihak karena frasa 'antar kelompok' tidak seharusnya diterapkan pada organisasi profesi. Institute for Criminal Justice Reform (ICJR) menyatakan bahwa penyeteraan profesi dengan suku, agama, dan ras yang seharusnya diatur dalam Pasal 28 ayat (2) UU ITE bisa berbahaya bagi iklim demokrasi di Indonesia (ICJR, 2021).

Seperti disebutkan sebelumnya, beberapa pengadilan telah "memperluas" penafsiran kata "kelompok" dengan mencakup pendukung Capres dan Cawapres, lembaga negara (misalnya, Komisi Pemilihan Umum)⁵⁶, bahkan bangsa Indonesia.⁵⁷ Salah satu alasannya adalah interpretasi inkonstitusional dari Mahkamah Konstitusi pada tahun 2017,⁵⁸ yang menyatakan bahwa "antar kelompok" mengacu pada realitas sosiologis keberadaan "kelompok lain" di luar suku, agama, dan ras. Meskipun fleksibilitas tersebut dapat digunakan untuk melindungi kelompok-kelompok yang tidak disebutkan secara eksplisit, seperti LGBT dan penyandang disabilitas, fleksibilitas tersebut hanya sering digunakan untuk melindungi mereka yang memiliki modal sosial yang cukup besar.

Mirip dengan penjelasan tentang pelaksanaan ketentuan mengenai pencemaran nama baik, beberapa kasus ujaran kebencian yang telah dijelaskan menunjukkan permasalahan yang menarik, dikarenakan pada

⁵⁵ Lihat Putusan Pengadilan Negeri Denpasar No. 72/Pid.Sus/2020/PT.Dps (14 Januari 2021)

⁵⁶ Lihat e.g., Putusan Pengadilan Negeri Jakarta Selatan No. 366/Pid.Sus/2019/PN.JKT.SEL (15 Agustus 2019) 84–85.

⁵⁷ Lihat Putusan Pengadilan Negeri Jayapura No. 16/Pid.Sus/2020/PN.Jap (29 April 2020) 38–39.

⁵⁸ Putusan Mahkamah Konstitusi Nomor 76/PUU-XV/2017 tentang Pengujian Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana telah diubah dengan Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (27 Maret 2018) 66–68

kasus-kasus tersebut, pengertian sewenang-wenang tentang “kelompok” sebagaimana diatur dalam Pasal 28 ayat (2) UU ITE telah diperparah. Dengan bagaimana terminologi tersebut dapat ditafsirkan secara luas, alih-alih melindungi masyarakat, penggunaan UU ITE yang sewenang-wenang membatasi kebebasan berbicara masyarakat di media sosial.

Moderasi Konten Berbasis Teknologi dan Otomatis

Pengguna media sosial dan algoritma media sosial keduanya berada dalam hubungan yang sama. Di satu sisi, platform-platform dan algoritma mereka mengetahui lebih dalam mengenai pengguna daripada diri mereka sendiri (Reviglio dan Agosti, 2019). Di sisi lain, terselubung oleh apa yang disebut “abstraksi perangkat lunak” (software abstraction) di mana detail internal sistem perangkat lunak sengaja disembunyikan, pengguna media sosial hampir tidak mengetahui apa-apa tentang entitas yang mengatur konten media sosial mereka (Zulli et al. 2020). Hal tersebut sama sekali tidak terkait dengan batasan kemampuan teknis – seperti tidak mengetahui cara coding – tetapi juga disebabkan oleh kode sumber yang tertutup (closed-source code), kekayaan intelektual, dan pembatasan hak cipta.

Pengguna media sosial tahu sangat sedikit dan bahkan tidak sama sekali mengenai bagaimana dan di mana data mereka disimpan, siapa yang memiliki akses ke data tersebut, bagaimana mereka dipindahkan ke seluruh jaringan, dan kesimpulan-kesimpulan apa yang dapat mereka ambil dari pola interaksi mereka di media sosial. Para pengguna berada di bawah pengawasan panopticon. Akan tetapi mata ini tidak hanya melihat tetapi juga pada akhirnya mendikte. Karen Yeung (2017) menulis tentang bagaimana platform menggunakan big data dan algoritma untuk mengendalikan tidak hanya feed media sosial tetapi juga perilaku pengguna. Yeung (2017, p. 120) berpendapat bahwa arsitektur platform-platform media sosial merupakan contoh penting dari sebuah dorongan: arsitektur yang dirancang secara sengaja yang mengubah perilaku orang dengan cara yang dapat diprediksi tanpa benar-benar membatasi opsi-opsi lainnya.

Reviglio dan Agosti (2020) pun melihat bagaimana, ketika ditempatkan dalam konteks algoritma media sosial, pilihan desain yang biasanya terpilih adalah yang berupaya membujuk—atau secara diam-diam memanipulasi—pengguna demi engagement dan monetisasi. Mereka berpendapat bahwa kondisi penelitian algoritma saat ini tidak dapat dilakukan baik ex-post atau ex-ante karena teknologi yang terus berubah dan sifat platform yang protektif. Sebagai saran, Reviglio dan Agosti mengusulkan audit algoritma platform untuk membawa lebih banyak partisipasi publik dan menjamin pengawasan yang demokratis. Hal ini dapat dilakukan dengan cara Mastodon dan Fediverse, yaitu, membuat kode dengan sistem open-source atau memberikan akses kepada lembaga masyarakat dan organisasi masyarakat sipil untuk membantu mengembangkan, mengawasi, dan mengembangkan laporan-laporan algoritma (Zulli et al., 2020).

Gagasan serupa dikemukakan oleh rekomendasi kebijakan Electronic Frontier Foundation (2020) untuk European Union's Digital Services Act. Selain audit, mereka juga menyarankan agar platform dapat lebih transparan ketika menggunakan algoritma sebagai proses pengambilan keputusan otomatis untuk memoderasi konten. Transparansi dapat berupa penandaan (flagging) pada langkah mana algoritma telah digunakan, menjelaskan logika di balik proses otomatis tersebut, dan menjelaskan bagaimana pengguna dapat menentang keputusan tersebut.

Membuat algoritma lebih transparan juga dapat mengubah persepsi publik untuk mengungkap proses-proses otomatis. Hal tersebut dilakukan dengan mengarusutamakan gagasan bahwa algoritma dan AI bukan sihir dan merupakan teknologi buatan manusia yang dapat memperburuk masalah-masalah yang ada ke dalam ranah daring. Salah satu contohnya adalah hambatan bahasa, yang mengakibatkan kurang atau lebihnya penandaan konten. (Barrett, 2020; Young, n.d.). Kesulitan ini terutama terlihat bagi mereka yang tinggal di negara-negara non-anglophone, terlebih lagi untuk negara-negara dengan banyak dialek daerah. Walau terdapat seorang moderator untuk meninjau keputusan yang dibuat oleh

mesin, sebuah konten akan dinilai berdasarkan keputusan yang dibuat oleh mesin terlebih dahulu, dan jumlah pekerja moderator konten juga terbatas, apalagi mereka dengan kondisi kerja dan kehidupan yang layak.

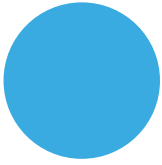
Algoritma-algoritma juga bias karena manusia yang memoderatori sistem tersebut juga berpotensi bias. Program merupakan produk sosio-teknis yang dilatih dengan database yang dipilih dan dianggap relevan oleh persepsi dan pengetahuan penciptanya. Tanpa pertimbangan yang peka terhadap perbedaan sosial, budaya, dan politik di dunia, moderasi konten otomatis pasti akan menghasilkan keputusan yang tidak efektif—bahkan berbahaya. Misalnya, ribuan video yang mendokumentasikan kekejaman selama Perang Saudara Suriah telah dihapus dari YouTube, dan postingan solidaritas dengan Palestina ditandai sebagai hasutan untuk melakukan kekerasan oleh Facebook (Gorwa et al., 2020; York dan Greene, 2021).

Platform mungkin berpendapat bahwa mereka telah memasukkan pengetahuan lokal untuk mengkontekstualisasikan kebijakan moderasi konten mereka. Di negara-negara non-anglophone, mesin dilatih untuk belajar dari kinerja moderator manusia lokalnya. Namun, peran moderator manusia ini seringkali terbatas pada penerapan pedoman yang ada. Mereka juga tidak serta merta dituntut untuk memiliki atau dibekali dengan kesadaran sosial-politik dan budaya tentang isu-isu sensitif. Mereka juga belum tentu memiliki sarana untuk memberikan masukan atas praktik dan proses moderasi konten (Ahmad, 2018). Singkatnya, partisipasi dari moderator manusia tidak memiliki makna. Standar-standar yang dibuat oleh kantor pusat yang mungkin tidak sensitif terhadap konteks lokal pun akan tetap ditegakkan. Mesin AI akan menerapkan praktik serupa dengan bias moderator manusia yang sama. Hal tersebut mungkin tidak efisien untuk mengatasi dan memerangi misinformasi dan disinformasi yang sedang berlangsung di komunitas non-bahasa Inggris. Sebuah laporan yang dibuat oleh Avaaz—sebuah kelompok keamanan online—menemukan bahwa Facebook gagal memproses 80% laporan tentang kesalahan informasi COVID-19 dalam bahasa Spanyol dibandingkan dengan hanya 27% dalam bahasa Inggris (Valencia, 2021). Di

Indonesia sendiri, tidak ada informasi yang tersedia tentang bagaimana platform mempekerjakan moderator ini atau berapa banyak moderator lokal yang dipekerjakan.

Merylina Lim dan Ghasad Alrasheed (2021) menyarankan langkah pertama dalam mengatasi bias algoritmik adalah dengan mengenali bahwa bias tersebut memang nyata. Ini mungkin cukup mudah untuk dikatakan, akan tetapi para pengembang sering kali mengabaikan kode performa dan menguniversalkan pelatihan mereka. Kemudian, komitmen yang kuat adalah suatu keharusan, seperti diversifikasi dataset yang digunakan untuk pelatihan atau bahkan melibatkan pengembang dari negara-negara yang terabaikan. Dalam hubungannya dengan mendorong lebih banyaknya transparansi dan pengawasan, bias yang telah dipelajari algoritma haruslah dihilangkan—begitu juga dengan bias yang telah dipelajari para pengembang.





BAB III

Kekhawatiran



Berdasarkan tren-tren yang digambarkan dalam bab sebelumnya, bab ini akan mengkaji berbagai isu yang muncul mengenai regulasi konten berbahaya dan ilegal. Bagian ini kemudian akan mengungkap wilayah abu-abu dari berbagai hal terkait konten berbahaya seperti ujaran kebencian, misinformasi, disinformasi, pencemaran nama baik, serta adanya kesenjangan antara mekanisme *self-regulatory* platform dengan peraturan dan kebijakan pemerintah yang berlaku. Pembahasan ini juga akan mencakup dampak hal-hal tersebut di masyarakat, termasuk dampak terhadap komunitas-komunitas yang rentan terhadap ketentuan hukum yang tidak adil.

Mengatur Wilayah Abu-abu: Ujaran Kebencian, Misinformasi dan Disinformasi, serta Pencemaran Nama Baik

Kebanyakan dari standar-standar internasional mengatur konten berbahaya dalam bentuk hukum lunak (*soft law*) yang tidak memiliki kekuatan mengikat. Namun, instrumen-instrumen hukum yang ada telah menguraikan apa saja yang harus dilakukan suatu negara dalam menyikapi berbagai hal terkait konten berbahaya di internet, antara lain ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik.

Standar-standar internasional yang mengatur ujaran kebencian dan pencemaran nama baik dapat ditemukan dalam laporan berjudul **Laporan Pelapor Khusus mengenai pemajuan dan perlindungan hak kebebasan berpendapat dan berekspresi terhadap Ujaran Kebencian Online (A/74/486)**. Laporan tersebut menegaskan kembali bahwa negara dapat membatasi kebebasan berekspresi sesuai dengan Pasal 19(3) ICCPR, yang mengharuskan semua peraturan untuk disediakan oleh hukum dan diperlukan untuk menghormati hak atau reputasi orang lain atau melindungi keamanan nasional, ketertiban umum, kesehatan masyarakat, atau moral. Terlebih lagi, peraturan yang dibuat oleh negara harus dilaksanakan secara ketat dan diawasi secara transparan. Instrumen yang sama pun memberikan sebuah contoh peraturan domestik yang tidak

mendefinisikan istilah-istilah kunci tetapi memberlakukan denda yang signifikan perusahaan yang gagal mematuhi ketentuannya. Peraturan tersebut kemudian dianggap kabur, dan definisi yang tidak jelas dianggap tidak konsisten dengan hukum hak asasi manusia internasional.

Sehubungan dengan disinformasi dan misinformasi, **Laporan Pelapor Khusus mengenai pemajuan dan perlindungan hak atas kebebasan berpendapat dan berekspresi (A/HRC/47/25)** lebih lanjut menetapkan bahwa undang-undang yang tidak jelas yang memberikan diskresi berlebihan dapat menyebabkan pengambilan keputusan yang sewenang-wenang dan tidak sesuai dengan Pasal 19(3) ICCPR. Instrumen khusus juga menyatakan bahwa kriminalisasi harus dilakukan hanya untuk kasus-kasus yang paling serius.

Berkaca pada apa yang telah diatur dalam beberapa standar internasional yang berkaitan dengan konten berbahaya, peraturan Indonesia masih jauh dari sesuai dengan standar. Peraturan-peraturan yang ada di Indonesia mengandung beberapa istilah yang memiliki multitafsir. Hal ini tercermin dari beberapa ketentuan dalam UU ITE terkait pencemaran nama baik, misinformasi, dan ujaran kebencian. Konstitusionalitas ketentuan ini juga telah beberapa kali diuji melalui pengajuan uji materiil ke Mahkamah Konstitusi. Namun, pengajuan-pengajuan yang telah dilakukan tidak menimbulkan hasil yang berarti. Dalam kasus-kasus pencemaran nama baik, meskipun Menkominfo, Kejaksaan Agung, dan Kapolri telah mengeluarkan pedoman pelaksanaan UU ITE (Keputusan Bersama), masih terdapat ambiguitas karena pedoman tersebut tidak merinci apakah suatu lembaga dapat dianggap sebagai objek pencemaran nama baik. Terkait dengan implementasi peraturan dan kebijakan yang ada, masih terjadi kriminalisasi terhadap warga sipil dalam penanganan kasus pencemaran nama baik, misinformasi, dan ujaran kebencian. Menurut SAFEnet, sepanjang tahun 2020, pejabat publik menduduki posisi teratas sebagai pelapor kriminalisasi hak kebebasan berpendapat di ruang digital dengan 47 kasus (SAFEnet, 2020). Hal tersebut menggambarkan bahwa peraturan dan kebijakan yang ada

cenderung digunakan untuk penguasa untuk membungkam kritik dan menimbulkan persekusi yang tidak adil.

Sebaliknya, standar internasional yang ada mengatur bahwa pengaturan mengenai ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik harus memenuhi asas legalitas yang mensyaratkan ruang lingkup, makna, dan akibat hukum yang jelas, tepat, dan umum. Lebih lanjut ditegaskan kembali bahwa undang-undang yang tidak jelas yang memberikan diskresi yang berlebihan dapat menyebabkan pengambilan keputusan yang sewenang-wenang dan tidak sesuai dengan Pasal 19(3) ICCPR dan oleh karena itu, tidak memenuhi standar internasional yang disyaratkan.

Selain itu, peraturan yang ada juga lebih mengutamakan pendekatan kriminal dibandingkan dengan metode penyelesaian lainnya (lihat Tabel 3). Hal ini tentunya tidak sejalan dengan standar internasional yang membedakan antara konten yang boleh dikriminalisasi dan konten yang jatuh diluar cakupan hukum pidana. Selain itu, semua Pelapor Khusus PBB tentang ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik menegaskan bahwa kriminalisasi harus dilakukan hanya untuk kasus-kasus yang paling serius. Oleh karena itu, kasus-kasus lain harus ditangani dengan berbagai pendekatan yang berbeda.

Terakhir, dapat disimpulkan bahwa mengatur ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik di Indonesia sama saja dengan mengatur wilayah abu-abu. Bukan hanya karena peraturan-peraturan yang ada saling terkait dan masih kabur, namun transparansi implementasi peraturan juga sering dipertanyakan. Selain itu, peraturan tersebut juga tidak berpihak pada minoritas, dan ada kecenderungan digunakan oleh para penguasa.

Kesenjangan antara Mekanisme Regulasi Mandiri Platform dengan Regulasi di Indonesia

Sudah menjadi fakta yang diketahui banyak orang bahwa konten diatur tidak hanya oleh negara namun juga oleh platform. Terdapat mekanisme-*self-regulatory*—atau kebijakan platform dan metode moderasi konten—yang terkadang tidak diatur oleh pemerintah domestik. Tanpa disadari, kita mungkin telah menjumpai beberapa mekanisme-mekanisme *self-regulatory* ini. Salah satunya adalah “Syarat dan ketentuan” atau “persyaratan layanan” yang sering kita setuju tanpa berpikir saat pertama kali membuat akun media sosial. Alat lain, yaitu ‘pedoman komunitas’, juga dapat diakses publik. Logika dari pedoman komunitas sederhana: setiap saat pengguna melanggar aturan yang telah dibuat dengan hati-hati, mereka dapat menghadapi konsekuensi *in-platform*, seperti penangguhan akun atau penghapusan konten.

Pada kenyataannya, menegakkan mekanisme pengaturan mandiri untuk konten—lebih dikenal sebagai moderasi konten—dapat menimbulkan beberapa tantangan bagi platform. Pertama, sulit bagi sebuah platform untuk menerapkan satu standar yang berlaku untuk konten secara global (Gillespie, 2018). Terkadang standar ini mungkin terlalu luas untuk dipahami oleh moderator berbasis teknologi. Misalnya, Facebook pernah menghapus konten pendidikan tentang menyusui karena salah mengartikan postingan tersebut sebagai pornografi yang ‘melanggar’ pedoman komunitasnya. (Gillespie, 2018). Di lain waktu, standar yang dibuat di markas platform ini mungkin dapat hilang dalam terjemahan, karena banyak bahasa-bahasa yang perlu ditafsirkan secara kontekstual sebelum dimoderasi (Wilson & Land, 2021, p. 1060, p.1064). Namun, ada dugaan bahwa investasi platform dalam memoderasi konten dalam bahasa non-Inggris sangat kekurangan dana. Menurut laporan Frances Haugen, seorang pelapor yang sebelumnya bekerja di Facebook, Facebook menghabiskan 87% dari dananya untuk memerangi misinformasi dalam bahasa Inggris; sementara itu, hanya 9% dari penggunanya yang benar-benar berbahasa Inggris (Valencia, 2021).

Facebook juga mengakui bahwa mereka terbiasa menyebarkan informasi yang memicu genosida di Myanmar dan gagal memoderasi konten ujaran kebencian yang memicu kejahatan perang (Stevenson, 2018).

Tantangan signifikan lainnya bagi platform-platform dalam meregulasi konten adalah dalam mengakomodasi hukum domestik dalam kebijakan moderasi konten mereka. Platform dan negara dapat memiliki perbedaan dalam pendekatan moderasi konten masing-masing. Di Indonesia, perbedaan-perbedaan ini bermula dari klasifikasi konten berbahaya (lihat tabel 2). Perbedaan cakupan konten berbahaya antara platform dan regulasi domestik Indonesia menyebabkan perbedaan dalam mekanisme penanganannya. Sebuah platform hanya akan menghapus konten-konten yang melanggar pedoman komunitasnya, namun sebagian besar akan menggunakan cara moderasi lain—yang dapat dianggap lebih halus—ketika konten tidak melanggar pedoman tetapi dapat dianggap berbahaya. Misalnya, platform hanya dapat menggunakan penandaan, pelabelan, penurunan setingkat, dan demonetisasi konten berbahaya (Audrine & Setiawan, 2021). Sebaliknya, Indonesia memperlakukan semua konten berbahaya—seperti yang telah diklasifikasikan dalam berbagai peraturan—sebagai ilegal. Dengan demikian, penghapusan konten adalah satu-satunya metode yang diakui secara hukum untuk memoderasi konten 'ilegal' di Indonesia berdasarkan peraturan yang ada.

Dalam sebuah buku putih yang baru-baru ini diterbitkan oleh Facebook (Bickert, 2020), para penulis mengungkapkan ketidakpuasan mereka tentang bagaimana definisi-definisi yang tidak jelas menghambat proses moderasi dan bahwa undang-undang tentang kewajiban penerbit juga dapat memiliki konsekuensi negatif seperti penyensoran berlebihan dan swasensor.

Tabel 2. Perbandingan Klasifikasi Konten dalam Peraturan Indonesia dan Pedoman Komunitas Platform⁵⁹

Klasifikasi Konten dalam Peraturan Indonesia	Facebook	Instagram	Twitter	YouTube	TIKTOK	WhatsApp	LINE
konten yang bertentangan dengan kepatutan (termasuk pornografi dan pornografi anak)	X	X	X	X	X	X	X
konten perjudian		X			X	X	
konten fitnah dan/atau pencemaran nama baik						X	
konten pemerasan dan/atau ancaman	X	X	X	X	X	X	
informasi palsu dan menyesatkan yang mengakibatkan kerugian konsumen dalam Transaksi Elektronik	X			X		X	X
informasi yang ditujukan untuk menimbulkan kebencian atau pertikaian terhadap individu dan/atau kelompok masyarakat tertentu berdasarkan suku, agama, ras, dan antargolongan	X	X	X	X	X	X	
konten yang mengandung ancaman kekerasan atau ketakutan yang ditujukan kepada individu	X	X	X	X	X		
Konten yang mengandung unsur kekerasan terhadap anak	X	X		X	X		X
Konten yang mengandung unsur pelanggaran kekayaan intelektual	X	X	X	X	X	X	X

⁵⁹ Dalam perbandingan ini, (x) menunjukkan jenis konten yang diatur sebagai 'ilegal' oleh pemerintah Indonesia dan diklasifikasikan sebagai 'berbahaya' oleh platform – karenanya memerlukan mekanisme moderasi konten untuk diberlakukan terhadap konten tersebut.

Klasifikasi Konten dalam Peraturan Indonesia	Facebook	Instagram	Twitter	YouTube	TIKTOK	WhatsApp	LINE
Konten yang mengandung unsur terorisme dan/atau radikalisme, separatisme dan/atau organisasi berbahaya yang dilarang	X	X	X	X	X	X	
Konten yang mengandung unsur pelanggaran perdagangan barang dan jasa melalui sistem elektronik	X	X				X	
Konten yang mengandung unsur pelanggaran keamanan informasi	X	X	X	X	X	X	X
Konten yang mengandung unsur pelanggaran	X					X	
Konten yang mengandung unsur pelanggaran di bidang kesehatan	X	X				X	
Konten yang mengandung unsur pelanggaran pengawasan obat dan makanan	X	X				X	
Konten yang mengganggu masyarakat dan ketertiban umum (misalnya, informasi dan/atau data yang dipalsukan)							X

Sumber: disusun oleh para penulis, 2021.

Platform-platform terutama lebih berhati-hati dalam mengatur konten ucapan, seperti dalam menangani misinformasi atau berita palsu. Twitter, misalnya, secara eksplisit menyatakan bahwa mereka "...melayani percakapan publik" dan ingin mempermudah pengguna untuk "...membuat keputusan yang tepat", (Roth & Pickles, 2020) kemudian memutuskan untuk melabeli postingan-postingan yang berisi informasi

menyesatkan atau menghapus konten tersebut berdasarkan tingkat bahaya yang dapat disebabkan oleh informasi tersebut. Facebook (sekarang Meta) memberlakukan strategi yang terdiri dari tiga langkah untuk melawan misinformasi atau berita palsu, yaitu dengan melakukan demonetisasi, menurunkan peringkat, lalu akhirnya menghapus konten atau halaman yang membagikan fakta yang dipalsukan (Meta, 2018). Meskipun penghapusan konten atau penangguhan akun adalah sebuah opsi, platform-platform biasanya hanya menggunakannya setelah tahap penilaian oleh moderator platform manusia atau mesin AI telah dilakukan. Proses penilaian yang dilakukan oleh platform ini dapat ditingkatkan agar lebih efisien dan tepat dalam menangani konten berbahaya. Namun, kami menemukan bahwa proses hukum khusus dalam mengatur konten ini belum ada dalam peraturan Indonesia (Audrine & Setiawan, 2021).

Audrine & Setiawan (2021) berpendapat bahwa kurangnya proses hukum yang semestinya meninggalkan platform-platform dengan dua pilihan ekstrem dalam memoderasi konten jika pemerintah mengajukan permintaan: untuk menyimpan atau menghapus konten yang dilaporkan. Jika mereka gagal untuk mematuhi, platform bertanggung jawab atas sanksi administratif mulai dari denda hingga penghentian akses. Jangka waktu singkat yang dinyatakan dalam Peraturan Menkominfo 5/20 dan ketakutan akan sanksi dapat memaksa platform untuk menghapus konten-konten tanpa penilaian yang cermat dan proses hukum yang semestinya. Selain itu, platform-platform tidak memiliki hak atau media untuk mengajukan banding atas keputusan yang dibuat oleh pemerintah. Dalam banyak hal, mirip dengan NetzDg, Peraturan Menkominfo 5/20 membentuk moderasi konten Indonesia dengan pendekatan hukuman. Pendekatan ini mungkin bukan pendekatan yang paling ideal untuk mekanisme moderasi konten yang ingin menegakkan asas-asas hak asasi manusia, terutama kebebasan berekspresi. Terlebih lagi, kurangnya proses hukum yang semestinya terjadi tanpa adanya mekanisme yang mengamankan transparansi, baik kepada pemerintah maupun platform. Akibatnya, baik pemerintah maupun platform hanya merilis informasi tentang jumlah dan jenis konten yang telah mereka moderasi, bukan

laporan transparansi yang berarti, yang menjelaskan cara mereka memoderasi konten daring.

Mekanisme hukuman dan jangka waktu yang singkat untuk memutuskan dan bertindak dapat mendorong platform untuk mengalihkan sumber daya mereka, menghentikan moderasi proaktif, dan hanya menunggu laporan (Bickert, 2020, pp. 13-14). Selain itu, platform juga dapat memprioritaskan peninjauan posting yang dilaporkan karena mereka lebih dekat dengan tenggat waktu 24 atau 4 jam daripada yang lebih lama (postingan yang tidak dilaporkan), meskipun konten itu juga dapat menyebabkan kerugian. Metrik lain yang dibahas—yang terlihat lebih disukai dalam buku putih yang telah disebutkan—adalah prevalensi atau jangkauan dan tingkat konten yang dilarang, berapa banyak yang melihat, menyukai, dan membagikannya. Dalam skenario ini, platform diberi insentif untuk lebih fokus pada satu konten berbahaya yang viral, dilihat oleh jutaan orang, daripada sepuluh atau dua puluh tetapi memiliki engagement yang sangat kecil. Oleh karena itu, metrik-metrik dapat diubah, meningkatkan sumber daya dan perhatian perusahaan pada area yang terukur dan pintasan yang tidak terukur. Hal ini menggemakan kekhawatiran yang diangkat oleh Alkiviadou (2019) bahwa siklus pemantauan berkorelasi dengan kinerja moderasi konten yang bergantung pada waktu.

Memang, Indonesia saat ini sedang berjuang dengan masalah ucapan hinaan. Di satu sisi, ujaran kebencian dan pencemaran nama baik seperti yang didefinisikan, ditafsirkan, dan diimplementasikan melalui KUHP dan UU ITE saat ini menghasilkan overkriminalisasi (Angendari, 2020; Hamid, 2019; Heryanto, 2021). Di sisi lain, peraturan-peraturan tersebut tidak mampu—atau lebih tepatnya pemerintah tidak mau—melindungi hak-hak minoritas, seperti minoritas agama dan gender (George, 2017). Namun, perusahaan media sosial, termasuk Facebook, tidak dikecualikan dari implementasi dokumen kebijakan yang tidak merata (Doctorow, 2021a;

Sinpeng & Martin, 2021). Faktanya, Electronic Frontier Foundation memiliki proyek yang didedikasikan seluruhnya untuk mendokumentasikan praktik-praktik ini (Trendacosta & York, 2019).

Tabel 3. Perbandingan Pendekatan Platform Media Sosial dan Negara dalam Menangani Konten Ilegal dan Berbahaya di Indonesia

	Negara	Platform Sosial Media
Metode	<ul style="list-style-type: none"> ● Pengadilan: pidana dan perdata ● Non-Pengadilan: alternatif penyelesaian sengketa, tindakan administratif (termasuk permintaan pemutusan akses dari masyarakat, polisi, jaksa, pengadilan, dan lembaga pemerintah lainnya) 	<ul style="list-style-type: none"> ● Penghapusan konten ● Penangguhan akun ● Penurunan peringkat ● Demonetisasi ● Penandaan ● Label dan peringatan
Subyek	<ul style="list-style-type: none"> ● Orang: mekanisme pengadilan, mekanisme non-pengadilan (terutama alternatif penyelesaian sengketa) ● Operator sistem elektronik: mekanisme non-pengadilan (terutama tindakan administratif), dan sampai batas tertentu dapat bertanggung jawab atas tindak 	<ul style="list-style-type: none"> ● Akun/pengguna: pembatasan (pelanggaran minimum) atau penangguhan (pelanggaran ganda)
Obyek	<ul style="list-style-type: none"> ● Pelanggaran: mekanisme pengadilan, mekanisme non-pengadilan (khususnya penyelesaian sengketa alternatif) ● Isi: mekanisme pengadilan (dalam hal perintah sementara atau permanen), tindakan administratif (permintaan penghapusan dari pemerintah) 	<ul style="list-style-type: none"> ● Konten: pengurangan visibilitas dan demonetisasi (non-pelanggaran tetapi dianggap berbahaya); penghapusan konten (pelanggaran atau sesuai

Sumber: disusun oleh para penulis, 2021

Dampak Regulasi Konten Terhadap Masyarakat

Terdapat beberapa cara bagaimana ujaran kebencian dan misinformasi dapat mempengaruhi masyarakat. Pada poin tertentu, keduanya dapat berbenturan dan saling terjalin dengan satu sama lain. Sebagai contoh, dis/misinformasi yang meliputi Pemilu tahun 2019 yang sebagian besar disebabkan oleh sentimen Anti-Cina (Temby, 2019).

Regulasi dari ujaran di daring mempengaruhi berbagai komunitas dengan cara yang berbeda-beda, beberapa diantaranya secara berat-sebelah lebih berpengaruh daripada komunitas lainnya. Dalam kasus-kasus tertentu, regulasi dapat digunakan bagi mereka yang berkuasa untuk membungkam suatu komunitas (seperti halnya politisi menggunakan regulasi pencemaran nama baik untuk mempersekusi para aktivis dan jurnalis), komunitas lainnya mungkin lebih tidak terlindungi di mayantara akibat identitas mereka (komunitas LGBTQ+), dan untuk komunitas lain, kasus ujaran kebencian yang dilakukan di ranah daring yang menyerang mereka kian meningkat (agama minoritas).

→ **Jurnalis**

Beberapa komentator telah mengingatkan bagaimana rezim regulasi konten saat ini dapat menyebabkan efek detrimental terhadap masyarakat Indonesia. Dalam dua kasus yang telah disebutkan di atas menggarisbawahi bagaimana tidak hanya masyarakat awam yang dapat terpengaruh, tetapi juga para jurnalis yang seharusnya terlindungi dengan hak-hak pers. Anton Muhajir (2019) yang menulis untuk Remotivi mengingatkan bahwa bagaimana hukum saat ini dapat memberikan “chilling effect” atau “efek pendingin” dimana masyarakat sipil dan jurnalis keduanya menjadi takut dan menyensor diri mereka karena takut dianggap bertanggung jawab atas ujaran mereka di ranah daring—yang kemungkinan valid. Tidak hanya takut akan polisi dan pejabat pemerintah, tetapi juga masyarakat lainnya. UU ITE telah sering digunakan untuk mempersekusi jurnalis yang dianggap mengancam. Sebagai contoh, jurnalis dan pembuat film dokumenter Dandhy Dwi Laksono dinamakan sebagai tersangka melanggar UU ITE akibat kicauannya di Twitter mengenai konflik di Papua pada tahun 2019 (Freiheit, 2021). Muhammad Asrul divonis 3 bulan penjara akibat ‘melanggar’ pasal 45 ayat 1 dan pasal 27 ayat 3 tentang ujaran kebencian setelah ia menulis artikel mengenai korupsi yang dilakukan oleh walikota Palopo yang diterbitkan di media daring beritanews.com (The Finery Report, 2021).

Selain itu, AJI (Aliansi Jurnalis Independen) telah mencatat 15 kasus mengenai kriminalisasi jurnalis selama tiga tahun terakhir (CNN Indonesia, 2021b). Di tahun 2021 saja, AJI menjelaskan bahwa Dewan Pers Indonesia menerima 44 laporan dari polisi mengenai dugaan pelanggaran UU ITE oleh para jurnalis (Ningtyas et al., 2021). Walaupun tidak semuanya dibawa ke Pengadilan, AJI dan Dewan Pers khawatir bahwa tren penggunaan UU ITE dan KUHP terhadap para jurnalis tidak menunjukkan penurunan. Pada tahun 2015, Kominfo Indonesia merilis pernyataan yang mengafirmasi bahwa para jurnalis yang memproduksi karya jurnalisme yang sesuai dengan UU Pers dan kode etik jurnalistik tidak perlu khawatir (Yura, 2015). Kominfo bersikeras bahwa UU ITE memiliki ketentuan untuk melindungi para jurnalis. Namun sebaliknya, berbagai jurnalis tengah menjadi korban walaupun mereka telah bekerja sesuai dengan UU Pers, metode jurnalistik, dan kode etik. Termasuk juga Asrul yang masih dipenjarakan walaupun hakim dan Dewan Pers menilai karyanya sebagai karya jurnalistik yang kredibel. Menurut UU Pers, subjek dari artikel-artikel berita berhak untuk mendapatkan hak untuk membalas, dan kanal berita tersebut bertanggung jawab untuk menghapus atau merubah artikel jika terdapat kesalahan fakta. Namun penggunaan UU ITE yang begitu sering oleh pejabat pemerintah dan masyarakat umum telah menjadikannya suatu norma.

→ Organisasi Masyarakat Sipil

Selain jurnalis, aktivis seringkali harus menyeimbangkan ujaran mereka walaupun mereka menggunakan bentuk-bentuk yang sah untuk mengevaluasi dan mengkritik. Dalam misi mereka untuk membantu mereka yang telah disalahkan oleh UU ITE, para aktivis dan anggota Organisasi Masyarakat Sipil (Ormas) sendiri harus berhati-hati kepada pejabat pemerintah yang menggunakan UU ITE untuk membungkam kritik dan penegak hukum yang menawarkan diskresi mereka kepada mereka yang berkuasa. Sebagai contoh, Haris Azhar, mantan Koordinator Komisi untuk Orang Hilang dan Korban Tindak Kekerasan (KontraS), dan Faida

Maulidiyanti, Koordinator Kontras saat ini, telah dilaporkan oleh Menteri Koordinator Maritim and Investasi Luhut Binsar Pandjaitan atas tindakan pencemaran nama baik dengan menggunakan UU ITE Pasal 27 ayat 3 (Maaruf, 2021). Kasus ini berawal dari video Youtube yang ditayangkan di kanal Haris Azhar yang berjudul "ADA LORD LUHUT DIBALIK RELASI EKONOMI-OPS MILITER INTAN JAYA!!JENDERAL BIN JUGA ADA!! NgeHAMtam". Video ini mendiskusikan sebuah laporan yang ditulis oleh beberapa Ormas yang menduga keterlibatan Pandjaitan dalam operasi pertambangan di Intan Jaya, Papua, Indonesia (Tim Advokasi Bersihkan Indonesia, 2021).

Fatia, Haris, dan para organisasi yang mengkompilasikan laporan tersebut berasumsi bahwa karena posisi Pandjaitan sebagai pejabat publik, mantan pejabat militer, dan jika ditambahkan dengan okupasi militer yang luar biasa di Intan Jaya kemudian menyebabkan suatu konflik kepentingan. Menteri Koordinator kemudian mengeluarkan surat panggilan terhadap Haris dan Fatia, mengatakan bahwa apa yang mereka katakan merupakan suatu fitnah dan berupaya untuk mendakwa mereka dengan pencemaran nama baik. Akan tetapi, Nurkholis, pengacara Haris berpendapat bahwa pihak Pandjaitan belum menjelaskan bagian mana dari video dan laporan tersebut yang mereka anggap tidak benar (Taher, 2021). Lebih jauh lagi, Asfinawati, pengacara Fatia, menyatakan bahwa Haris dan Fatia menggunakan hak mereka untuk berekspresi bebas dengan memikirkan kepentingan publik dan menargetkan untuk membahas posisi Pandjaitan sebagai pejabat publik dan bukan sebagai individu. Sehingga pengacara Fatia berpendapat bahwa jika penelitian yang disitasi oleh Faia dan Haris mengandung informasi tidak benar, isu tersebut seharusnya diselesaikan dengan klarifikasi dan bukan dengan surat panggilan. Hingga saat ini, kasus tersebut masih berlanjut, dan polisi sedang berproses dari surat panggilan menjadi investigasi formal (CNN Indonesia, 2022).

➔ Masyarakat

Ketika jurnalis dan aktivis dapat dibungkam, maka posisi masyarakat umum menjadi lebih rentan. Sebagai contoh, WP dan AS, dua rakyat sipil yang terjerat oleh UU ITE Pasal 28 ayat 2 tentang ujaran kebencian (SAFE-net, 2021). WP, seorang buruh di Riau, ditangkap oleh polisi atas tuduhan penghinaan terhadap Presiden Indonesia Joko Widodo. Ia mengunggah sebuah meme di Facebook yang berisi wajah Presiden Jokowi dan kalimat menyindir yang mengindikasikan ketidakpuasan beliau dengan kebijakan pemerintah Indonesia mengenai COVID-19. Ia kemudian didakwa oleh pihak berwajib dengan UU ITE Pasal 28 ayat 2 tentang ujaran kebencian dan, menurut polisi, komentar tersebut dapat berpotensi untuk memicu konflik horizontal (CNN Indonesia, 2020). Sementara itu, AS juga ditangkap dan didakwa dengan UU ITE Pasal 27 ayat 3 tentang pencemaran nama baik dan penghinaan terhadap Pemerintah Kota Semarang. Menurut polisi, laporan atas AS datang langsung dari Pemerintah Kota Semarang yang merasa tersinggung akibat postingan AS di Facebook yang mengandung umpatan dan kritik terhadap penutupan jalan sebagai kebijakan untuk menghentikan penyebaran pandemi COVID-19 (Rahmayadi, 2020). Keduanya diekspresikan dengan maksud untuk mengkritik, salah satunya untuk menghibur, dan keduanya dikatakan dalam konteks posisi pemerintah dalam menghadapi pandemi COVID-19. Dalam waktu krisis seperti halnya pandemi, kedua ekspresi ini merupakan sebuah bentuk katarsis ketidakpuasan mereka dengan keadaan di negara mereka dan kebijakan yang diambil pemerintah mereka. Ekspresi mereka ini sangatlah masuk akal dan sama sekali tidak dimaksud sebagai suatu ujaran kebencian.

➔ Komunitas Terpinggirkan

Ariel Heryanto (2021), akademisi media dan Profesor Studi Indonesia, Australia, menjelaskan bahwa daripada hukum yang melindungi pejabat pemerintah dan kelompok dan individu yang berkuasa lainnya dari kritik yang sah serta bentuk lainnya dan ujaran di ranah daring, harus ada ketentuan yang diperuntukkan untuk melindungi martabat masyarakat.

Melindungi mereka dari pejabat-pejabat sombong yang menyalahgunakan kekuasaan dan koneksi mereka. Hal ini berhubungan dengan fakta bahwa dalam rezim hak asasi manusia internasional, pencemaran nama baik bukan lagi dianggap sebagai tindakan yang masuk dalam hukum pidana. Namun masuk ke ranah hukum privat. Heryanto menganggap pemenjaraan akibat mengkritik pemerintah—mendefinisikannya sebagai pencemaran nama baik, memicu kebencian, atau memberikan misinformasi pada publik—merupakan suatu peninggalan jaman kolonial yang perlu dihapuskan.

Kumpulan dari UU ITE yang tidak sempurna, penegak hukum yang tidak akuntabel, serta pejabat-pejabat yang sombong merupakan alasan mengapa Electronic Frontier Foundation dan SAFEnet memperingatkan kita semua akan regulasi yang meminta platform untuk bertanggung jawab kepada pemerintah di masa depan (Rodriguez, 2021)—mengatakan bahwa Peraturan Menteri Komunikasi dan Informatika No. 5 sebagai “bentuk pelanggaran HAM”. Mereka menyarankan untuk mereformasi hukum-hukum yang relevan sebelum meminta pertanggungjawaban platform kepada pemerintah.

Menurut riset yang dilakukan oleh SAFEnet, berbagai pasal digunakan untuk mengkriminalisasi pengguna internet pada tahun 2020 (SAFEnet, 2021). Dalam 84 kasus yang disebutkan di atas, sebagian besar kasus menggunakan pasal-pasal problematis dari UU ITE, terutama Pasal 28 ayat 2 tentang ujaran kebencian dalam 27 kasus, Pasal 27 ayat 3 tentang pencemaran nama baik dalam 22 kasus, dan Pasal 28 ayat 1 tentang kerugian konsumen akibat informasi yang salah dalam 12 kasus. Peningkatan jumlah kriminalisasi terhadap kebebasan berekspresi tidak dapat dipisahkan dari bagaimana pemerintah menghadapi pandemi COVID-19 dan penerbitan UU Cipta Lapangan Kerja. Disitasi dari SAFEnet, berbagai kasus yang terjadi berkaitan dengan kedua isu tersebut.

Mengenai metode-metode non-pengadilan, Kominfo, dari laporan tahunan 2020, menyatakan bahwa jumlah penghentian akses ke konten ilegal di laman situs berjumlah 130,254 (Kementerian Komunikasi dan Informatika, 2021). Sementara, jumlah keseluruhan dari penghentian akses ke konten ilegal (dan membahayakan) di platform media sosial mencapai 183,434. Penghentian akses dilakukan dengan menggunakan Mesin AIS, yang menggunakan Kecerdasan Buatan untuk mencari konten ilegal di ruang digital dengan menggunakan metode pencarian berdasarkan kata kunci.

Dari total 130,254 konten internet ilegal di laman-laman situs, konten perjudian menjadi yang paling umum di angka 76,216, kemudian diikuti dengan 46,172 konten pornografi, 3,484 konten yang mengandung penipuan, 2,903 adalah konten yang melanggar hak cipta, 1,366 adalah konten ilegal yang dilaporkan oleh badan-badan sektor, 95 adalah konten yang melanggar keamanan informasi, 10 disinformasi, dan 8 konten yang mengandung terorisme, radikalisme, sentimen terhadap SARA, separatisme, organisasi berbahaya, dan lain sebagainya. Ditambah lagi, Twitter memiliki konten ilegal terbanyak dengan 165,698 konten; Facebook, Instagram, dan WhatsApp dengan 5,483 konten; pembagian dokumen dengan 1,272 konten; Google dan Youtube dengan 399 konten; Telegram dengan 225 konten, dan LINE dengan 1 konten.

Dalam aplikasinya, regulasi konten yang ada di regulasi Indonesia masih memiliki berbagai masalah yang mengkhawatirkan. Jika dibiarkan, terdapat kekhawatiran terjadinya kriminalisasi berlebihan, khususnya terhadap masyarakat sipil.

Dalam regulasi yang ada, beberapa terminologi memiliki interpretasi ganda. Hal ini terlihat dalam beberapa ketentuan di UU ITE terkait dengan pencemaran nama baik, misinformasi dalam transaksi elektronik, ujaran kebencian, dan penyadapan. Bahkan, konstitusionalitas dari ketentuan ini mengandung istilah-istilah ambigu yang telah diuji beberapa kali selama

lewat uji materi konstitusional ke Mahkamah Konstitusi.⁶⁰ Dari berbagai permintaan, hanya satu yang dikabulkan oleh Mahkamah Konstitusi mengenai penyadapan ilegal.

Dalam kasus-kasus mengenai pencemaran nama baik, pengadilan dalam Kasus Nomor 223/Pid.Sus/2018/PN Kbm, Pengadilan merferensikan Keputusan Mahkamah Agung No. 183 K/Pid/2010 yang memperluas objek dari pencemaran nama baik yang menyatakan bahwa suatu entitas hukum dapat menjadi objek pencemaran nama baik.⁶¹ Walaupun Kominfo, Jaksa Agung, dan Kepala Kepolisian Republik Indonesia mengeluarkan petunjuk pelaksanaan UU ITE (Keputusan Bersama) yang meregulasikan lebih jauh mengenai pencemaran nama baik, hal itu tidak secara spesifik apakah suatu institusi dapat menjadi objek pencemaran nama baik. Ambiguitas makna ini kemudian menjadi kekhawatiran lain dalam mengimplementasikan regulasi konten yang sudah ada saat ini.

Selain keberadaan interpretasi ganda atas istilah-istilah yang digunakan, regulasi yang sudah ada juga memprioritaskan kriminalisasi dibandingkan dengan metode penyelesaian perkara lainnya. Hal ini terlihat dari pemetaan klasifikasi konten daring yang didasari regulasi Indonesia yang ada pada Tabel 1. Dari berbagai jenis konten yang diatur, hanya 2 (dua) jenis konten yang membolehkan upaya administratif dalam penanganannya. Hal ini tidaklah sesuai dengan standar internasional, yang membedakan antara konten yang dapat dikriminalisasi dan yang tidak.

Selain dari masalah regulasi konten, negara dan platform media sosial seringkali memiliki pendapat yang berbeda mengenai konten yang dianggap ilegal dan membahayakan. Hal ini terlihat dalam berbagai definisi dari konten yang ilegal dan membahayakan yang tertera di regulasi negara dan petunjuk pemakaian media sosial. Media sosial menamakan konten

⁶⁰Lihat Putusan Mahkamah Konstitusi Nomor 50/PUU-VI/2008 (perkara pencemaran nama baik), Putusan Mahkamah Konstitusi Nomor 2/PUU-VII/2009 (perkara pencemaran nama baik), Putusan Mahkamah Konstitusi Nomor 52/PUU-XI/2013 (perkataan kebencian) kasus), Putusan Mahkamah Konstitusi No. 76/PUU-XV/2017 (kasus ujaran kebencian)

⁶¹Lihat Putusan Mahkamah Agung No. 183 K/Pid/2010 (20 Mei 2010) 15.

yang dapat ditandai atau dihapuskan sebagai “konten membahayakan”, negara menggunakan istilah “konten yang dilarang”. Hal ini kemudian menyebabkan pendekatan yang berbeda dalam menangani konten yang dimaksud. Selain itu, terdapat ketentuan-ketentuan dalam mengimplementasikan regulasi UU ITE yang tidak sesuai dengan petunjuk penggunaan platform media sosial. Sebagai contoh, regulasi yang ada hanya mengadopsi mekanisme penghapusan konten dalam menangani konten. Sementara, media sosial telah memiliki berbagai mekanisme untuk menangani konten seperti menandakan konten yang dianggap membahayakan.

Kurangnya kesepakatan suara atas definisi dari konten-konten ilegal dan membahayakan di kerangka hukum Indonesia dan platform media sosial dan ketidakjelasan masing-masing istilah (seperti moralitas, ketertiban umum, dll.) akan mempengaruhi perlindungan terhadap kebebasan menyatakan pendapat. Hal ini memperlihatkan ketidakpastian hukum, potensi untuk merugikan komunitas terpinggirkan seperti jurnalis, aktivis, dll. Seperti yang telah disebutkan sebelumnya, pemerintah dan pihak yang berkuasa adalah subjek yang seringkali menyalahgunakan regulasi yang sudah ada untuk membungkam masyarakat umum yang ingin menyuarakan opini mereka mengenai suatu kebijakan tertentu yang dikeluarkan oleh pemerintah. Namun, status-quo dari ketetapan ketentuan konten yang dianggap ilegal dan membahayakan telah menyebabkan represi terhadap kritik dan ketidakpuasan terhadap pemerintah yang ter ‘desentralisasi’, dimana tidak ada lagi kolaborasi nasional antara masyarakat umum dan pemerintah, namun menjadi diskresi penguasa untuk melindungi kepentingan mereka (Hamid, 2019).

● **Komunitas Terpinggirkan: LGBTQ+**

Sayangnya, dalam perlindungan terhadap hak komunitas terpinggirkan, platform media sosial tidaklah aman dan seringkali menjadi lingkungan yang agresif kepada lesbian, gay, biseksual, transgender, *queer*, dan lainnya (“LGBTQ+”) secara virtual. Konten

yang mereka pos di media sosial dimoderasi secara sewenang-wenang oleh platform media sosial yang akan mengikuti standar dan pedoman komunitas, yang kemudian membuat banyak akun LGBTQ+, postingan, dan iklan tematik dihapuskan akibat laporan dan/atau diblokir prosedur mereka. Sementara, konten homofobik, transfobia, dan seksis tidak tersentuh (EDRI, 2019). Komunitas LGBTQ+ pastinya telah menghadapi berbagai rintangan akibat adanya standar ganda.

Meskipun demikian, Indonesia tidak secara eksplisit mengkriminalisasi LGBTQ+ karena mengekspresikan identitas dan praktik seksual di media sosial dalam regulasi nasionalnya. Terdapat berbagai batasan, seperti apa yang diregulasikan di Pasal 4 ayat 1 dan apa yang dijelaskan dalam UU Pornografi yang melarang menciptakan, diseminasi, atau menayangkan pornografi yang mengandung hubungan seksual, termasuk seks antara lesbian dan seks homoseksual antara laki-laki. Namun, konstitusi Indonesia telah menetapkan hak bagi seluruh Indonesia, yang seharusnya juga termasuk mereka yang merupakan bagian dari komunitas LGBTQ+—walaupun hal ini tidak disebutkan secara eksplisit, untuk bebas dari tindakan diskriminatif.

Di sisi lain, terdapat berbagai kasus menyangkut diskriminasi dan pembatasan dari hak untuk berekspresi dan berpendapat. Sebagai contoh, untuk merayakan *pride month* (bulan kebanggaan—bulan untuk memperingatkan kebanggaan kaum LGBTQ+ atas perjuangan mereka untuk mendapatkan HAM), Persatuan Wartawan Indonesia untuk Keberagaman telah menciptakan sebuah ruang digital untuk solidaritas komunitas lewat *webinar* langsung di bulan Juni 2020 (New Mandala, 2020). Akan tetapi, harapan untuk menyampaikan informasi penting mengenai hak dan perlindungan LGBTQ+ ketika Youtube, sebagai respon kepada komplain yang diajukan penonton mereka mengenai betapa sensitifnya konten tersebut, menghapus webinar tersebut di tengah siaran. Lebih jauh lagi, aktivis *gay* Hartoyo menghadapi tantangan virtual dalam gerakannya. Ketika Jakarta

memasuki *lockdown* di bulan April, pameran luring yang merupakan bagian dari penggalangan dana program lewat butik SriKendes yang menjual baju-baju motif tradisional asli kemudian beralih ke pameran virtual. Hartoyo mulai menggunakan Facebook, dengan nama akun “Har Toyo” dan kemudian mulai menggunakan model penggalangan dana baru dengan melelang baju-baju, tas, dan sepatu bekas lewat Facebook Live. Namun, para pengguna melaporkan akunnya dan berakhir pada dua kali penangguhan akun Facebook-nya. Alasan mengapa Facebook Indonesia menangguhkan akun tersebut adalah akibat penggunaan kata *bencong-an* yang merupakan istilah peyoratif bahasa Indonesia yang bermakna sama dengan “*queer*”. Kata tersebut diklasifikasikan oleh algoritma Facebook sebagai ujaran kebencian (Norman Harsono, 2020). Terlebih lagi, terdapat berbagai tindakan yang dilakukan oleh Kominfo untuk membatasi banyaknya konten LGBTQ+ dengan memblokir laman-laman situs mereka (lihat Kementerian Komunikasi dan Informatika 2018a; Kementerian Komunikasi dan Informatika, 2016; Kementerian Komunikasi dan Informatika 2018b).

Sinpeng et al. (2021) melihat presensi tidak menyenangkan dari ujaran kebencian di kolom komentar tiga halaman komunitas LGBTQ+ di Facebook. Komentar negatif yang berada di masing-masing halaman Facebook tersebut mewakili setidaknya sepertiga dari seluruh komentar publik. Terdapat peningkatan angka dalam komentar yang mengandung unsur kebencian dan ujaran kebencian itu sendiri sejak Menteri Riset, Teknologi, dan Pendidikan Tinggi melontarkan saran untuk melarang mahasiswa LGBTQ+ dari kampus untuk melindungi norma dan moralitas di Indonesia. Lewat wawancara dengan para pengelola masing-masing halaman komunitas LGBTQ+ tersebut, komentar dan ujaran kebencian tersebut tidak hanya lazim di Facebook, tetapi juga di berbagai platform lain seperti Twitter, Youtube, dan Instagram—bahkan para pengelola juga diserang dan diganggu lewat kanal pesan pribadi (Sinpeng et al., 2021).

Penerimaan dan toleransi dari LGBTQ+ di Indonesia tidak dapat dipisahkan dari sikap sosio-kultural dan nilai-nilai religius. Penerimaan berarti mereka dapat berpartisipasi dalam segala kehidupan keluarga dan sosial tanpa pengecualian, sementara toleransi lebih sering diekspresikan secara terpaksa atau akibat keharusan (UNDP & USAID, 2014). Agama telah menjadi bagian penting dalam mengkonstruksikan norma sosial dan hukum di Indonesia. Secara garis besar, seluruh agama yang diterima di Indonesia tidaklah setuju dengan praktik LGBTQ+ (Andina, 2016). Sesuai dengan hal tersebut, mayoritas nilai sosio-kultural dan norma di Indonesia juga telah menolak keberadaan LGBTQ+—walaupun telah ada beberapa perubahan dalam konteks kekakuan, dan budaya yang semakin terbuka terhadap keberagaman.

● **Komunitas Terpinggirkan: Agama Minoritas**

Selain komunitas LGBTQ+, salah satu kelompok terpinggirkan yang telah menjadi target ujaran kebencian di Indonesia adalah komunitas Ahmadiyah—sebuah sekte Islam yang kerap dianggap menodai Islam yang merupakan agama mayoritas di Indonesia (CSIS Indonesia, 2021; Solikhin, 2016). Berdasarkan data dari Center for Strategic and International Studies (CSIS) Indonesia, jumlah ujaran kebencian terhadap kelompok Ahmadiyah sebagai kelompok agama minoritas di media sosial, terutama Twitter, dari 1 Januari 2019 hingga 31 Juli 2021 telah mencapai angka yang mengkhawatirkan. Pada tahun 2008, Sobri Lubis, salah satu pemimpin dari Front Pembela Islam (FPI), muncul dalam tayangan video di Youtube, berkhotbah sambil berteriak “Bunuh! Bunuh! Bunuh! Bunuh para Ahmadi!” (George, 2017). Alisa Wahid, Koordinator Umum dari Jaringan Gusdurian menyatakan bahwa provokasi dari ujaran kebencian di media sosial seringkali digunakan untuk melakukan tindakan kekerasan di dunia nyata (Apriyani, 2017).

Di tahun 2016, keributan di Tanjung Balai, Sumatera Utara, dipicu oleh pesan berantai yang beredar lewat berita dari mulut-ke-mulut

dan media sosial (The Jakarta Post, 2016). Pesan tersebut telah diputar, terdistorsi dari sebuah percakapan yang menyebutkan bagaimana pengeras suara masjid semakin tahun semakin keras akibat tuduhan bahwa seorang non-Muslim berniat untuk melarang adzan (Pusat Studi Agama dan Demokrasi Paramadina, 2018). Gerombolan umat Muslim yang kesal dengan sentimen ini kemudian menghancurkan kuil umat Buddha dan tiga pagoda. Menghancurkan juga peralatan sembahyang, meja, kursi, mobil, motor, patung, dan menjarah rumah-rumah ibadah dalam prosesnya.

Mirip dengan apa yang terjadi di Singkil, Aceh dan Tolikara, Papua. Konflik antar-agama, yang kemudian menyebabkan sebuah gereja dibakar di Singkil, dipercaya terjadi akibat ujaran kebencian yang tersebar di media sosial (Lubabah, 2015). Konten anti-Syiah juga prevalen dalam media sosial Indonesia (Azali, 2017). Lebih banyak kejadian dapat ditemukan mengenai interseksi antara agama dan ketimpangan kuasa yang terjadi akibat perbedaan antara kaum mayoritas dan minoritas (mentalitas 'kita dan mereka').

Islamofobia—ujaran kebencian dan ketakutan terhadap Islam—di media sosial telah meningkat di Indonesia. Hal ini terjadi akibat pengaruh eksternal, visi orientalis mengenai Islam dari barat, dan tiga faktor internal (Kastolani, 2021). Pertama, reaksi terhadap khotbah intoleran dari tokoh masyarakat religius yang fundamentalis. Kedua, bentuk kebebasan berekspresi. Dan ketiga, akumulasi dari identitas politik dan polarisasi religius warganet Indonesia.

Panjaitan (2016) menjelaskan bahwa pemerintah yang permisif serta tingkat literasi yang rendah dari pengguna internet di Indonesia saling terjalin untuk menciptakan lingkungan yang tidak aman maupun damai bagi agama minoritas di media sosial. Walaupun Indonesia telah memiliki regulasi terhadap ujaran kebencian, baik daring maupun luring, tindakan ini tetap sering terjadi. Panjaitan menawarkan dua kemungkinan mengapa hal ini terjadi. Pertama, hal

ini terjadi karena masyarakat umum tidak menyadari adanya regulasi ini. Kedua, hal ini terjadi akibat kurangnya penegakan regulasi dari pemerintah akan hukum ini, terutama untuk melindungi agama minoritas.

Akan tetapi melarang ujaran kebencian, menurut Panjaitan (2015), merupakan hal yang dilematis. Secara historis, pemerintah Indonesia—walaupun sudah melewati transisi dari otoritarianisme ke demokrasi—telah menggunakan hukum untuk membungkam dan mengkriminalisasi siapapun yang mengkritik pemerintah. Menurut Panjaitan, solusi harus dipikirkan di level masyarakat, adalah dengan meningkatkan literasi digital seperti saring sebelum membagikan konten.





BAB IV

Kesimpulan dan Rekomendasi



Kesimpulan

Dengan meningkatnya jumlah pengguna media sosial dari tahun ke tahun, upaya untuk meregulasi konten ilegal berbahaya di internet merupakan hal yang sangat penting. Hal ini juga akibat fenomena dimana peningkatan sirkulasi konten tidak beriringan dengan angka literasi digital masyarakat yang ada di Indonesia. Melihat kompleksitas dan ramifikasi dari isu ini, penelitian ini menunjukkan bahwa upaya untuk meregulasi konten ilegal dan membahayakan bukan hanya tanggung jawab pemerintah, tetapi juga sektor swasta seperti platform media sosial. Berdasarkan analisis dari kerangka hukum nasional, swa-regulasi dari platform media sosial, serta implementasinya, penelitian ini menggarisbawahi tiga temuan penting mengenai regulasi konten ilegal dan membahayakan serta implementasinya di Indonesia seperti yang akan dijelaskan di bawah ini.

➔ **Klasifikasi Konten dan Definisi**

Seperti yang dapat dilihat pada Tabel 1, temuan penelitian menunjukkan bahwa regulasi dan kebijakan Indonesia tidaklah membedakan antara konten yang 'ilegal' dan konten yang 'membahayakan'. Hampir seluruh konten membahayakan diklasifikasikan sebagai ilegal dan kemudian dianggap sebagai tindak pidana. Sehingga, publikasi atau diseminasi dari konten tersebut menjadi berada dibawah sanksi kriminal. Sehingga dapat disimpulkan bahwa regulasi konten di Indonesia belum sesuai dengan standar internasional yang telah mencoba untuk membedakan antara konten 'membahayakan' dan konten 'ilegal'.

Dalam ranah internasional, bahkan instrumen hukum mengikat yang membedakan antara konten membahayakan dan ilegal juga jarang ditemukan. Seringkali, perbedaan yang ditemukan berada dalam bentuk 'lunak' yang belum mengikat secara hukum. Namun, instrumen masih dapat digunakan sebagai referensi. Hal ini diperkuat dengan pentingnya membedakan konten 'membahayakan' dan

konten 'ilegal' dalam meregulasikan konten di media sosial karena dapat terjadi perbedaan dalam penanganan dan kewajiban untuk peduli pada konten ilegal dan membahayakan dalam praktik antara negara dan platform media sosial.

Sehingga, terdapat urgensi untuk mengadopsi standar internasional dalam mengkategorisasikan konten ilegal dan konten membahayakan. Standar tersebut kemudian diharapkan dapat menjadi tolok ukur bagi para negara untuk dapat menangani konten ilegal dan membahayakan dengan lebih teratur.

Kemudian, terdapat ketidaksamaan yang mencolok antara klasifikasi konten membahayakan yang digunakan oleh pemerintah dan platform. Ketidaksamaan ini kemudian dapat menyebabkan perbedaan dalam menangani konten daring, sehingga mengurangi efisiensi moderasi konten platform. Platform juga merupakan subjek dari regulasi konten domestik, dimana mereka menghadapi dilema antara mengkompromikan mekanisme regulasi-diri mereka atau dihukum akibat tidak mengikuti ketentuan.

➔ Mekanisme Penanganan

Dari Tabel 2, kami menemukan bahwa masing-masing stakeholder memiliki pendekatan yang berbeda dalam menangani konten ilegal dan membahayakan. Lewat regulasi dan kebijakan, negara memiliki dua cara dalam menangani konten ilegal dan membahayakan di Indonesia, yaitu jalur hukum dan jalur non-hukum. Jalur hukum termasuk menangani kasus-kasus lewat pengadilan pidana dan sipil. Di sisi lain, jalur non-hukum termasuk penyelesaian perkara alternatif dan tindakan administratif seperti penghapusan konten dan platform, denda administratif, dll.

Di Indonesia, mekanisme penanganan dapat berbeda secara signifikan dengan negara. Metode moderasi konten yang secara hukum diakui di Indonesia adalah penghapusan konten dan memblokir akses. Sementara platform memiliki beberapa metode

untuk digunakan, yaitu: *downranking*, mencabut keuntungan finansial dari suatu konten, penandaan, pelabelan, dan peringatan, menghapus konten, dan menanggukkan akun. Platform juga sangat berhati-hati dalam konten ujaran. Mereka dapat menolak untuk menghapus konten berita palsu tetapi dapat mengurangi visibilitas berita tersebut. Dengan Permenkominfo No. 5 tahun 2020, platform dapat terkena hukuman jika mereka menolak untuk menghapus konten, hal ini dapat memperlihatkan pendekatan penghukuman dalam regulasi konten di Indonesia. Hal ini dapat mengancam hak masyarakat karena dapat berdampak penyensoran, penyensoran diri, serta tindakan-tindakan lain yang dapat melanggar hak-hak fundamental.

→ Masalah lainnya

Walaupun regulasi dan kebijakan negara—juga pedoman komunitas platform media sosial—telah ada dan telah memberikan berbagai mekanisme untuk menangani konten ilegal dan membahayakan, masalah-masalah untuk secara efektif meregulasikan konten ilegal dan membahayakan di internet masih tersisa.

1

Pertama, regulasi dan kebijakan negara—terutama yang terkait dengan ujaran kebencian, misinformasi, disinformasi, dan pencemaran nama baik—tidaklah sesuai dengan standar internasional yang ada. Walaupun kejelasan merupakan syarat utama dalam membuat suatu regulasi menurut standar internasional, terdapat masalah-masalah normatif seperti halnya istilah ambigu yang dapat diinterpretasikan secara luas dalam regulasi dan kebijakan negara, serta tidak adanya perbedaan eksplisit antara ilegal dan membahayakan seperti yang telah dijelaskan sebelumnya.

2

Kedua, terdapat permasalahan implementasi terkait dengan subjektivitas dan interpretasi yang berbeda-beda oleh pihak yang berwajib, termasuk pejabat pemerintah, polisi, jaksa, pengadilan, dan bahkan masyarakat sendiri. Masalah ini terlihat dalam

perbedaan substansi dalam berbagai kebijakan nasional yang dikeluarkan dan berbagai konsiderasi dari pengadilan distrik yang seperti yang telah dijelaskan di atas. Sehingga, isu-isu yang telah ada memiliki efek detrimental kepada komunitas-komunitas terpinggirkan seperti halnya komunitas LGBTQ+ dan minoritas religius.

3

Ketiga, terdapat masalah dalam hubungan antara negara dan platform media sosial dalam menangani konten ilegal dan membahayakan. Dalam kasus-kasus tertentu, masih terdapat ketidaksetujuan antara negara dan platform media sosial dalam menangani konten tersebut. Hal ini terlihat dari perbedaan-perbedaan dari konten ilegal dan membahayakan dalam regulasi negara dan pedoman penggunaan media sosial.

4

Keempat, platform media sosial harus meningkatkan mekanisme regulasi-diri mereka. Platform harus berinvestasi lebih kepada moderasi konten berbahasa non-Inggris dan berkomitmen untuk meningkatkan praktik moderasi konten, data menunjukkan bahwa inisiatif ini masih belum optimal, terutama untuk memerangi misinformasi di ranah daring. Moderator konten lokal masih belum memiliki peran yang berarti dalam memoderasi konten lokal, dan laporan transparansi yang sudah ada tidak memiliki informasi proses moderasi.

5

Kelima, belum ada ketentuan yang mengharuskan transparansi yang berarti dalam moderasi konten di Indonesia. Walaupun baik pemerintah dan platform merilis informasi mengenai jenis dan jumlah dari konten yang telah mereka moderasikan, mereka tidak memberikan informasi detail mengenai bagaimana mereka memutuskan suatu konten sebagai membahayakan. Kondisi ini, ditambah dengan tidak adanya proses hukum, praktik moderasi konten di Indonesia dapat berpotensi untuk mengancam kebebasan berekspresi di ranah daring.

Namun demikian, dengan waktu dan sumber daya yang singkat, penelitian sebagian besar dilakukan dengan menganalisis data sekunder. Sehingga, banyak hal mengenai implementasi regulasi konten masih dapat dikembangkan dan dieksplorasi secara empiris, sebagai contoh, mengenai dampak terhadap komunitas yang termarjinalkan. Lebih jauh lagi, penelitian ini akan menjadi dasar dari aktivitas Fase II dari proyek SM4P dengan mengatur suatu koalisi *multi-stakeholder* nasional di Indonesia, mengembangkan mekanisme pemantauan daring untuk konten berpotensi ilegal atau membahayakan, serta mengorganisir konsultasi dengan komunitas terdampak-konflik untuk memahami dampak dari konten membahayakan terhadap dinamika konflik dan meningkatnya aksi intoleransi terhadap agama, gender, dan seksual minoritas di Indonesia. Penelitian ini juga diharapkan untuk mendorong diskusi dan penelitian lebih lanjut mengenai dampak regulasi konten bagi masyarakat dan kelompok minoritas termarginalisasi di Indonesia. Terakhir, lewat berbagai temuan dan rekomendasi yang dijelaskan dalam riset ini serta aktivitas yang melibatkan banyak *stakeholders* selama proyek ini berlangsung, keseluruhan proyek SM4P diharapkan dapat mengumpulkan berbagai pemikiran dan kekhawatiran dari berbagai pemangku kepentingan untuk menstimulasi perubahan-perubahan lebih baik dalam meregulasi konten dan menangani konten berbahaya untuk pemerintah Indonesia dan *platform* media sosial.

Rekomendasi

Hasil dari pemetaan dan analisis hukum memperlihatkan bahwa masih banyak ruang untuk memperbaiki dan meningkatkan efektivitas regulasi konten ilegal dan membahayakan di Indonesia. Seperti yang telah disebutkan sebelumnya, dapat dilihat bahwa langkah menuju reformasi bukanlah merupakan tanggung jawab tunggal dari negara. Sehingga, komitmen dari sektor swasta, seperti platform media sosial, juga sangat diperlukan.

Rekomendasi 1:



Merevisi UU ITE dan regulasi implementasinya

UU ITE berfungsi sebagai dasar hukum utama dalam meregulasi ruang siber di Indonesia. Sehingga, UU ITE juga memiliki regulasi pengimplementasian, seperti Peraturan Pemerintah No. 71 tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik dan Permenkominfo No. 5 tahun 2020, yang juga telah dijelaskan di pemetaan hukum. Namun, dalam implementasinya, beberapa hal harus diperbaiki, baik dalam UU ITE dan regulasi pengimplementasiannya. Berikut adalah beberapa diantaranya:



Mereformasi ulang regulasi klasifikasi konten

Penelitian ini merekomendasikan negara untuk mereformasi ulang jenis konten yang ada di UU ITE dan regulasi pengimplementasiannya. Pengadilan seringkali menggunakan pendekatan konvensional yang kemudian menyebabkan kriminalisasi berlebihan. Reformasi ulang ini dapat dilakukan dengan mengadopsi standar internasional. Saat ini, beberapa instrumen hukum internasional telah memberikan dasar untuk membedakan antara konten ilegal dan membahayakan—atau setidaknya memetakan konten yang dapat dikriminalisasi dan yang tidak. Sebagai contoh, Laporan dari Pelapor Khusus mengenai promosi dan perlindungan hak untuk kebebasan beropini dan berekspresi (A/66/290) membedakan tiga tipe konten, yaitu: (a) ekspresi yang merupakan bentuk pelanggaran dari hukum internasional dan dapat dihukum secara pidana; (b) ekspresi yang tidak dapat dihukum secara pidana tapi dapat dilarang dan mendapatkan gugatan sipil; dan (c) ekspresi yang tidak dapat diberi sanksi secara pidana maupun sipil, tetapi tetap mengkhawatirkan dalam aspek toleransi, kesopanan, dan rasa saling hormat antar manusia.

● **Mendefinisikan ulang istilah-istilah yang terkait dengan konten ilegal dan membahayakan**

Penelitian ini merekomendasikan Negara untuk mendefinisikan ulang istilah-istilah yang terkait dengan konten ilegal dan berbahaya. Dalam kasus-kasus tertentu, hakim, polisi, dan jaksa menggunakan istilah yang memiliki makna luas dalam menangani kasus-kasus yang berkenaan dengan ilegal dan membahayakan. Sebagai contoh, kasus-kasus mis/disinformasi ditangani dengan pasal-pasal yang terkait dengan pencemaran nama baik. Dalam konteks daring, terdapat istilah-istilah baru yang memiliki definisi sendiri dan tidak dapat diinterpretasikan dengan istilah yang sudah ada sebelumnya, seperti misinformasi dan disinformasi.

● **Mereformasi mekanisme penanganan konten**

Penelitian ini merekomendasikan Negara untuk mereformasi mekanisme penanganan konten dalam UU ITE dan regulasi pengimplementasiannya saat ini. Pembaruan ini termasuk beberapa hal. *Pertama*, regulasi akan mengakomodir berbagai jenis mekanisme penanganan konten. Saat ini, regulasi yang sudah ada hanya mengadopsi mekanisme penghapusan konten. Sementara, platform media sosial sudah memiliki berbagai mekanisme penanganan konten, seperti halnya menandai konten yang dianggap membahayakan. *Kedua*, regulasi akan menambahkan waktu bagi platform media sosial untuk menangani konten yang telah ditandai. Regulasi saat ini sangatlah kaku terhadap waktu, walaupun platform media sosial harus memeriksa konten terlebih dahulu sebelum menghapusnya. *Ketiga*, regulasi akan mengurangi keringanan sanksi bagi platform media sosial. Tingginya angka sanksi yang diregulasikan dalam regulasi saat ini membebani para platform dan membuat mereka menciptakan mekanisme untuk menghindari sanksi. *Keempat*, regulasi akan menambahkan mekanisme banding untuk platform media sosial. Kelima, mereformasi ulang sanksi untuk kasus-kasus yang berkaitan dengan konten yang membahayakan.

Seperti yang telah dijelaskan sebelumnya, sebagian besar sanksi terhadap kasus-kasus konten membahayakan adalah sanksi pidana. Akan tetapi, upaya hukum lewat sanksi pidana terhadap konten membahayakan harusnya menjadi pilihan terakhir. Pendekatan komprehensif lainnya lewat edukasi dan cara-cara teknologis lainnya dapat dilakukan.

Rekomendasi 2:

- ➔ Mengharmonisasi hukum dan regulasi yang berkaitan dengan konten ilegal dan membahayakan

Penelitian ini merekomendasikan negara untuk mengharmonisasi hukum dan regulasi yang berkaitan dengan konten ilegal dan membahayakan. Harmonisasi tersebut diproyeksikan dapat mengurangi kemungkinan perbedaan interpretasi dan ketumpang-tindihan antar regulasi. Rekomendasi ini tidak hanya dilakukan dalam regulasi yang mengelola daring tetapi juga ke regulasi lain yang dapat bersimpangan dengan isu konten ilegal dan membahayakan.

Rekomendasi 3:

- ➔ Menyetarakan persepsi dari makna ketentuan atas tindakan yang dilarang dalam UU ITE dan UU lain yang terkait

Penelitian ini merekomendasikan Negara untuk memberikan suatu persepsi umum dari makna ketentuan atas tindakan yang dilarang dalam UU ITE dan UU lain yang terkait. Dalam kasus-kasus tertentu, suatu tindakan dapat diinterpretasikan berbeda oleh polisi, jaksa, dan pengadilan. Sehingga, harus ada persepsi seragam mengenai ketentuan dari tindakan apa yang dilarang. Hal ini dapat dicapai dengan menciptakan keputusan bersama atau instrumen hukum lainnya. Terlebih lagi, dalam ujaran kebencian, mis/disinformasi, dan kasus-kasus pencemaran nama baik, polisi, jaksa, dan pengadilan juga dapat mengadopsi tes ambang batas enam-bagian (*six-part threshold test*) yang didasari oleh Rabat Plan of Action dalam menilai apakah suatu konten membahayakan, yaitu (a)

konteks, (b) pembicara, (c) niat, (d) konten dan bentuk, (e) cakupan dari tindakan ujaran, dan (f) kemungkinan. Tes ini dapat diinkorporasikan ke dalam regulasi seperti halnya UU ITE.

Rekomendasi 4:

- Meningkatkan kerjasama antara Negara dan platform media sosial dalam menangani konten ilegal dan membahayakan

Penelitian ini merekomendasikan negara dan platform media sosial dapat meningkatkan kerjasama mereka, dimana dalam menangani konten membahayakan, harus ada tanggung jawab bersama antara negara dan platform media sosial. Dalam kasus ini, harus dipastikan bahwa negara dan platform media sosial bergerak ke arah yang sama dalam menangani konten ilegal dan membahayakan yang beredar di daring, terutama konten ujaran kebencian, mis/disinformasi, dan pencemaran nama baik yang tengah meningkat selama beberapa tahun terakhir. Sehingga, berbagai diskusi dan pertemuan dengan banyak pemangku kepentingan harus didukung.

Rekomendasi 5:

- Meningkatkan transparansi dalam memoderasi konten

Penelitian ini merekomendasikan negara dan platform media sosial memastikan transparansi yang berarti dalam mengimplementasikan regulasi konten mereka. Transparansi yang berarti berarti bahwa mereka tidak berhenti menciptakan laporan yang berbasis tindakan yang telah dilakukan (seperti angka konten yang telah dimoderasi) tetapi juga menginformasikan masyarakat tentang bagaimana proses moderasi konten itu dilakukan.



Bibliography

- Ahmad, S. (2018). "It's Just the Job": Investigating the Influence of Culture in India's Commercial Content Moderation Industry.
- Alexandra, L., Satria, A., Suryahudaya, E.G., & Krisetya, B. (2021). The National Hate Speech Dashboard. CSIS Indonesia. Retrieved from: <https://dashboard.csis.or.id/hatespeech/>
- Alkiviadou, N. (2019). Hate speech on social media networks: Towards a regulatory framework? *Information & Communications Technology Law*, 28(1), 19–35. <https://doi.org/10.1080/13600834.2018.1494417>
- Amnesty International Indonesia, et al. (2021), *Kertas Kebijakan: Catatan dan Desakan Masyarakat Sipil atas Revisi UU ITE*.
- Andina. (2016). Faktor Psikososial dalam Interaksi Masyarakat dengan Gerakan LGBT di Indonesia, 7(2), 173–185. <http://jurnal.dpr.go.id/index.php/aspirasi/article/view/1288/709>
- Angendari, D.A.D. (2020, November 30). Definisi "ujaran kebencian" di Indonesia terlalu luas, gampang dimanfaatkan. *The Conversation*. <http://theconversation.com/definisi-ujaran-kebencian-di-indonesia-terlalu-luas-gampang-dimanfaatkan-150743>
- Anindyajati, T. (2021). Limitation of the Right to Freedom of Speech on The Indonesian Constitutional Court Consideration. *Indonesian Law Journal*, 14(1), 19–36. <https://doi.org/10.33331/ilj.v14i1.45>
- Antara and Kukuh S. Wibowo (2021). Terapkan Surat Edaran Kapolri Soal UU ITE, Kasus Novel Baswedan akan Dimediasi <https://nasional.tempo.co/read/1435856/terapkan-surat-edaran-kapolri-soal-uu-ite-kasus-novel-baswedan-akan-dimediasi> (diakses pada 2 Februari 2022)
- Apriyani, R. (2017). Gusdurian: Provokasi Ujaran Kebencian di Medsos Dorong Kekerasan di Dunia Nyata. *KBR*. https://kbr.id/nasional/02-2017/gusdurian__provokasi_ujaran_kebencian_di_medsos_dorong_kekerasan_di_dunia_nyata/88912.html
- Aschoff, N. (2020, Mei 29). Social Media Companies Can't Be Trusted to Protect Our Democracy. *Jacobin*. <https://jacobinmag.com/2020/05/donald-trump-twitter-social-media-companies>
- ASEAN Declaration to Prevent and Combat Cybercrime 2017.
- ASEAN Framework and Joint Declaration to Minimize the Harmful Effects of Fake News 2018.

- Association for Progressive Communication (APC). (2018). Content Regulation in the Digital Age. Submission to the United Nations Special Rapporteur on the Right to Freedom of Opinion and Expression.
- Audrin, P & Setiawan, I. (2021). Impact of Indonesia's Content Moderation Regulation on Freedom of Expression. CIPS Policy Paper (38). <https://repository.cips-indonesia.org/media/347642-impact-of-indonesias-content-moderation-88ec21cc.pdf>
- Azali, K. (2017). Fake News and Increased Persecution in Indonesia. *ISEAS Perspective*, 61, 10.
- Banchik, A.V. (2021). Disappearing Acts: Content Moderation and Emergent Practices to Preserve at-risk Human Rights Related Content. *New Media & Society* 23(6), 1527-1544.
- Banko, M., MacKeen, B., and Ray, L. "A Unified Typology of Harmful Content", Proceedings of the Fourth Workshop on Online Abuse and Harms, 20 November 2020, <https://aclanthology.org/2020.alw-1.16.pdf>
- Barrett, P. (2020). Who Moderates the Social Media Giants? A Call to End Outsourcing (p. 32). NYU Stern Center for Business and Human Rights. https://bhr.stern.nyu.edu/tech-content-moderation-june-2020?_ga=2.264803983.1095278896.1633969975-1758350172.1632663756
- Benedek, W., & Kettemann, M. C. (2013). Freedom of Expression and the Internet. Council of Europe. https://book.coe.int/fr/attachment?id_attachment=709
- Bernie, M. (2020, Februari 3). Cara Murah Sikapi Berita: 2 Jurnalis Sulawesi Dijerat Pasal Karet. *tirto.id*. <https://tirto.id/cara-murahan-sikapi-berita-2-jurnalis-sulawesi-dijerat-pasal-karet-ewsR>
- Bickert, M. (2020). Charting a Way Forward on Online Content Regulation. Facebook. https://about.fb.com/wp-content/uploads/2020/02/Charting-A-Way-Forward_Online-Content-Regulation-White-Paper-1.pdf
- Branden, B., Davidse, S., and Smit, E. (2021). In between illegal and harmful: a look at the Community Guidelines and Terms of Use of online platforms in the light of the DSA proposal and the fundamental right to freedom of expression. DSA Observatory. <https://dsa-observatory.eu/2021/08/02/in-between-illegal-and-harmful-a-look-at-the-community-guidelines-and-terms-of-use-of-online-platforms-in-the-light-of-the-dsa-proposal-and-the-fundamental-right-to-freedom-of-expression-part-1-of-3/>

- Briantika, Adi (2020). "Lima Telegram Kapolri jadi Pedoman Tangani Pelanggaran PSBB", <https://tirto.id/lima-telegram-kapolri-jadi-pedoman-tangani-pelanggaran-psbb-eLk4> (diakses pada 2 Februari 2022)
- Budapest Convention on Cybercrime 2001.
- Burns, H. (2020). Online Harms: Freedom of Expression Remains Under Threat. Open Rights Group. <https://www.openrightsgroup.org/blog/online-harms-freedom-of-expression-remains-under-threat/>
- CNN Indonesia. (2020, April 8). Diduga Hina Jokowi soal Corona, Buruh di Kepri Ditangkap. <https://www.cnnindonesia.com/nasional/20200408192303-12-491818/diduga-hina-jokowi-soal-corona-buruh-di-kepri-ditangkap>
- CNN Indonesia. (2022, Januari 6). Kasus Luhut Vs Haris Azhar Naik Penyidikan di Polda Metro Jaya. CNN Indonesia. <https://www.cnnindonesia.com/nasional/20220106160339-12-743459/kasus-luhut-vs-haris-azhar-naik-penyidikan-di-polda-metro-jaya>
- CNN Indonesia. (2021a). Google: Indonesia Jadi Negara Terbanyak Minta Hapus Konten. *teknologi*. Retrieved Februari 3, 2022, from <https://www.cnnindonesia.com/teknologi/20211025150358-185-712049/google-indonesia-jadi-negara-terbanyak-minta-hapus-konten>
- CNN Indonesia. (2021b). AJI: 3 Tahun Terakhir Ada 15 Jurnalis Media yang Dijerat UU ITE. <https://www.cnnindonesia.com/nasional/20211201151322-12-728525/aji-3-tahun-terakhir-ada-15-jurnalis-media-yang-dijerat-uu-ite> (diakses pada 7 Februari 2022)
- Cochrane, J. (2014, Juni 3). Embrace of Atheism Put an Indonesian in Prison. *The New York Times*. https://www.nytimes.com/2014/05/04/world/asia/indonesian-who-embraced-atheism-landed-in-prison.html?_r=0
- Commission of the European Communities. (1996). Illegal and harmful content on the Internet – Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. Brussels, 16 October 1996.
- Council of the European Union. (2014). EU Human Rights Guidelines on Freedom of Expression Online and Offline.
- CSIS Indonesia. (2021). CSIS National Hate Speech Dashboard. CSIS Indonesia. <https://dashboard.csis.or.id/hatespeech/>
- De Gregorio, G. (2020). Democratizing Online Content Moderation: A Constitutional Framework. *Computer Law & Security Review* 36.

- De Streef, A. (2020). Online Platforms' Moderation of Illegal Content Online. Study requested by the European Parliament's Committee on Internal Market and Consumer Protection.
- Debora, Y. "Daftar Pasal UU ITE yang Sering Menjerat Netizen di Medsos", <https://tirto.id/daftar-pasal-uu-ite-yang-sering-menjerat-netizen-di-medsos-gbdg>
- Detik (2021) "Para Penunggang UU ITE" [https://news.detik.com/x/detail/investigasi/20210301/Para-Penunggang-UU-ITE/\(diakses pada 16 Oktober 2021\)](https://news.detik.com/x/detail/investigasi/20210301/Para-Penunggang-UU-ITE/(diakses%20pada%2016%20Oktober%202021))
- Dimas Jarot Bayu, "Indonesia Negara yang Paling Banyak Blokir Iklan Daring." <https://databoks.katadata.co.id/datapublish/2021/03/03/indonesia-negara-yang-paling-banyak-blokir-iklan-daring> (diakses pada 30 November 2021)
- Doctorow, C. (2021a, Juli 16). Right or Left, You Should Be Worried About Big Tech Censorship. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2021/07/right-or-left-you-should-be-worried-about-big-tech-censorship>
- Doctorow, C. (2021b, Agustus 3). With Great Power Comes Great Responsibility: Platforms Want To Be Utilities, Self-Govern Like Empires. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2021/08/utilities-governed-empires>
- Deklarasi Universal Hak Asasi Manusia 1948.
- EDRI. (2019). The Digital Rights of LGBTQ+ People: When Technology Reinforces Societal Oppresions. European Digital Rights. <https://edri.org/our-work/the-digital-rights-lgbtq-technology-reinforces-societal-oppressions/> (diakses pada 10 Desember 2021)
- European Parliament. (2020). Report on Digital Services Act and fundamental rights issues.
- European Union Communication on Illegal and Harmful Content on the Internet 1996.
- Evandio, A. CIPS: Revisi UU ITE Diperlukan, Masih Banyak Pasal Multitafsir. Bisnis. <https://kabar24.bisnis.com/read/20210908/15/1439555/cips-revisi-uu-ite-diperlukan-masih-banyak-pasal-multitafsir> (diakses pada 2 Desember 2021).

- Evans, D. G. (2009). Human Rights and State Fragility: Conceptual Foundations and Strategic Directions for State-Building. *Journal of Human Rights Practice*, 1(2), 181–207.
- Farras, B. (2019). Astaga! Facebook Paling “Bandel” Tak Ikuti Arahan Pemerintah. *CNBC Indonesia*. Retrieved Februari 3, 2022, from <https://www.cnbcindonesia.com/tech/20190514094646-37-72267/astaga-facebook-paling-bandel-tak-ikuti-arahan-pemerintah>
- Flew, T., Suzor, N., & Martin, F. (2019). Internet Regulation as Media Policy: Rethinking the Question of Digital Communication Platform Governance. *Journal of Digital Media Policy* 10(1), 33-50.
- Freiheit. (2021). Press Freedom: Indonesia Government Fears on Critics and Murals. *Friedrich Naumann Foundation.org*. <https://www.freiheit.org/indonesia/press-freedom-indonesia-government-fears-critics-and-murals> (diakses pada 7 Februari 2022)
- George, C. (2017). Indonesia: Demokrasi yang Diuji di Tengah Intoleransi. In *Pelintiran Kebencian: Rekayasa Ketersinggungan Agama dan Ancamannya bagi Demokrasi* (pp. 145–178). PUSAD Paramadina and IIS UGM. <https://www.paramadina-pusad.or.id/buku/pelintiran-kebencian-rekayasa-ketersinggungan-agama-dan-ancamannya-bagi-demokrasi/>
- Gorwa, R. (2019). The Governance Triangle: Conceptualising the Informal Regulation of Online Content. *Internet Policy Review* 8 (2), 1-22.
- Gorwa, R., Binns, R., & Katzenbach, K. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*.
- Government Regulation No. 71 of 2019 on Electronic System and Transaction Implementation.
- Gelber, K. (2019). Differentiating Hate Speech: A Systemic Discrimination Approach. *Critical Review of International Social and Political Philosophy*, 24 (4). Retrieved from: <https://doi.org/10.1080/13698230.2019.1576006>
- Hamid, U. (2019, November 20). UU ITE dan merosotnya kebebasan berekspresi individu di Indonesia. *The Conversation*. <http://theconversation.com/uu-ite-dan-merosotnya-kebebasan-berekspresi-individu-di-indonesia-126043>
- Hamid, U. (2019, November 25). Indonesia’s Information Law has Threatened Free Speech for More than a Decade. This Must Stop. *Theconversation.com*. <https://theconversation.com/indonesias-information-law-has-threatened-free-speech-for-more-than-a-decade-this-must-stop-127446> (diakses pada 3 desember 2021)

- Harian Jogja (2022). KontraS Beberkan Fakta Demokrasi Memburuk Begini Cara Penguasa Membungkam Masyarakat Sipil, <https://news.harianjogja.com/read/2022/01/06/500/1092729/kontras-beberkan-fakta-demokrasi-memburuk-begini-cara-penguasa-membungkam-masyarakat-sipil> (diakses pada 2 Februari 2022)
- Harsono, N. (2020). In Conservative Indonesia, This Gay Activist Braves Social Curbs to Help the Marginalized. *The Jakarta Post*. <https://www.thejakartapost.com/news/2020/08/25/in-conservative-indonesia-this-gay-activist-braves-social-curbs-to-help-the-marginalized.html> (diakses pada 10 Desember 2021)
- Hootsuite. (2021). Digital 2021, Global Overview Report. https://hootsuite.widen.net/s/zcdrtxwczn/digital2021_globalreport_en
- Heryanto, A. (2021, Juni 26). Gila Hormat. *Kompas.id*. <https://www.kompas.id/baca/opini/2021/06/26/gila-hormat-2/>
- Indonesian Institute of the Independent Judiciary. (2018). Penafsiran terhadap Pasal 156A KUHP tentang Penodaan Agama (Analisis Hukum dan Hak Asasi Manusia). Policy Brief.
- Institute for Criminal Justice Reform (ICJR). (2021, Januari 19). Putusan Banding Jerinx: Hakim Gagal Koreksi Pertimbangan yang dapat Berujung pada Malapetaka di Indonesia. <https://icjr.or.id/putusan-banding-jerinx-hakim-gagal-koreksi-pertimbangan-yang-dapat-berujung-pada-malapetaka-di-indonesia/> (diakses pada 6 Februari 2022)
- International Covenant on Civil and Political Rights (ICCPR) 1976.
- International Covenant on Economic, Social and Cultural Rights (ICESCR) 1976.
- International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) 1969.
- Jakarta Globe (2015) "Police to Name Suspects in Tolikara Arson Case: Chief", <https://jakartaglobe.id/news/police-name-suspects-tolikara-arson-case-chief/> (diakses pada 2 Februari 2022)
- Kastolani. (2020). Understanding the delivery of Islamophobic hate speech via social media in Indonesia. *Indonesian Journal of Islam and Muslim Societies*, 10(2), 247–270. <https://doi.org/10.18326/ijims.v10i2.247-270>
- Katadata. (2020). Puncak Penyebaran Hoaks Terjadi Menjelang Pilpres 2019: Jumlah Temuan Hoaks (2019). *Katadata.com*. <https://databoks.katadata.co.id/datapublish/2020/01/07/jelang-pemilu-hoaks-makin-berseliweran> (diakses pada 3 Desember 2021).

- Keen, C., Kramer, R., & France, A. (2020). The Pornographic State: The Changing Nature of State Regulation in Addressing Illegal and Harmful Content Online. *Media Culture and Society* 42(7-8), 1175-1192.
- Kementerian Kesehatan Republik Indonesia. (2019). Kemenkes Meminta Kemkominfo Blokir Iklan Rokok di Internet. P2PTM.Kemkes.go.id. <http://p2ptm.kemkes.go.id/kegiatan-p2ptm/pusat-/kemenkes-meminta-kemkominfo-blokir-iklan-rokok-di-internet> (diakses pada 16 Oktober 2021).
- Kementerian Komunikasi dan Informatika. (2016). Kemenkominfo: Stiker LGBT LINE Tidak Berlaku di Indonesia. *Kominfo.go.id*. https://kominfo.go.id/content/detail/6730/kemenkominfo-stiker-lgbt-line-tidak-berlaku-di-indonesia/0/sorotan_media (diakses pada 10 Desember 2021).
- Kementerian Komunikasi dan Informatika. (2017). Berantas Iklan Kesehatan Hoaks, Kemenkes Gandeng 7 Lembaga. *Kominfo.go.id*. https://kominfo.go.id/index.php/content/detail/12121/berantas-iklan-kesehatan-hoaks-kemenkes-gandeng-7-lembaga/0/sorotan_media (diakses pada 30 November 2021)
- Kementerian Komunikasi dan Informatika. (2018a). Ini Dia Upaya Gigih Menkominfo Babat Habis LGBT. *Kominfo.go.id*. https://www.kominfo.go.id/content/detail/12448/ini-dia-upaya-gigih-menkominfo-babat-habis-lgbt/0/sorotan_media (diakses pada 10 Desember 2021).
- Kementerian Komunikasi dan Informatika. (2018b). Penanganan Konten yang Melanggar Nilai dan Norma Sosial Budaya. *Kominfo.go.id*. https://kominfo.go.id/index.php/content/detail/12403/siaran-pers-no08hmkominfo012018-tentang-penanganan-konten-yang-melanggar-nilai-dan-norma-sosial-budaya/0/siaran_pers (diakses pada 10 Desember 2021)
- Kementerian Komunikasi dan Informatika. (2019) Siaran Pers No. 67/HM/KOMINFO/03/2019. *Kominfo.go.id*. https://www.kominfo.go.id/content/detail/17465/siaran-pers-no-67hmkominfo032019-tentang-kominfo-pantau-dan-siap-blokir-iklan-kampanye-di-platform-digital-selama-masa-tenang/0/siaran_pers (diakses pada 30 November 2021)

- Kementerian Komunikasi dan Informatika. (2021). Annual Report Ministry of Communications and Information [Kominfo.go.id](https://www.kominfo.go.id/content/detail/36485/annual-report-ministry-of-communication-and-information-technology/0/laporan_tahunan).
https://www.kominfo.go.id/content/detail/36485/annual-report-ministry-of-communication-and-information-technology/0/laporan_tahunan (diakses pada 3 desember 2021).
- Kementerian Komunikasi dan Informatika. (2021). Siaran Pers No. 143 / H M / K O M I N F O / 0 4 / 2 0 2 1 .
https://kominfo.go.id/content/detail/34136/siaran-pers-no-143hmkominfo042021-tentang-sejak-2018-kominfo-tangani-3640-ujaran-kebencian-berbasis-sara-di-ruang-digital/0/siaran_pers (diakses pada 3 desember 2021).
- Keputusan Bersama Menteri Komunikasi dan Informatika RI., Jaksa Agung RI, dan Kepala Kepolisian Negara RI No. 229 of 2021, No. 154 of 2021, No. KB/2/VI/2021.
- Keputusan Mahkamah Konstitusi No. 50/PUU-VI/2008 tentang Peninjauan Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945.
- Keputusan Mahkamah Konstitusi No. 1/PUU-XI/2013 tentang Pengujian Undang-Undang Nomor 1 Tahun 1946 tentang KUHP dan Undang-Undang Nomor 8 Tahun 1981 tentang Hukum Acara Pidana terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945.
- Keputusan Mahkamah Konstitusi No. 52/PUU-XI/2013 tentang Peninjauan Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik Terhadap Undang-Undang Dasar Negara Republik Indonesia Tahun 1945
- Keputusan Mahkamah Konstitusi No. 76/PUU-XV/2017 tentang Peninjauan Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana telah diubah dengan Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.
- Keputusan Mahkamah Konstitusi No. 48/PUU-VII/2010 tentang Peninjauan Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi
- Keputusan Mahkamah Konstitusi No. 82/PUU-XVIII/2020 tentang Peninjauan Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi

- Khan, I. (2021). Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression on disinformation and freedom of expression (A/HRC/47/25).
- Kitab Undang-Undang Hukum Pidana Republik Indonesia.
- Koalisi Tolak Kriminalisasi dan Rekayasa Kasus. (2020, April 23). [PERNYATAAN BERSAMA KOALISI TOLAK KRIMINALISASI DAN REKAYASA KASUS] Segera Lepaskan Rasio Patra, Hentikan Kriminalisasi, Ungkap Pelaku Peretasan! Institute for Criminal Justice Reform. <https://icjr.or.id/pernyataan-bersama-koalisi-tolak-kriminalisasi-dan-rekayasa-kasus-segera-lepaskan-rasio-patra-hentikan-kriminalisasi-ungkap-pelaku-peretasan/>
- KontraS (2021) Pemutakhiran Data Virtual Police, <https://kontras.org/2021/04/22/pemutakhiran-data-virtual-police/> (diakses pada 2 Februari 2022).
- Koran Tempo. (2020, Februari 10). Stop Kriminalisasi Sadli. Tempo. <https://kolom.tempo.co/read/1305497/stop-kriminalisasi-sadli>
- La Rue, F. (2011). Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (A/66/290).
- Lubabah, R. (2015, November 2). Edaran hate speech disebabkan kasus Tolikara dan Aceh Singkil. Merdeka.Com. <https://www.merdeka.com/peristiwa/edaran-hate-speech-disebabkan-kasus-tolikara-dan-aceh-singkil.html>
- Lundahl, O. (2020). Algorithmic meta-capital: Bourdieusian analysis of social power through algorithms in media consumption. *Information, Communication & Society*, 1-16. <https://doi.org/10.1080/1369118X.2020.1864006>
- Madiega, T. (2020, Mei). Reform of the EU liability regime for online intermediaries. In-depth Analysis of European Parliamentary Research Service.
- Mahkamah Konstitusi Republik Indonesia. (2010). Naskah Komprehensif Perubahan UUD 1945 Buku VIII. Sekretariat General and Registrar of the Constitutional Court.
- Mahkamah Konstitusi. (2020, Oktober 27). Petitioner of Pornography Law Strengthens Argument. MKRI.id. <https://en.mkri.id/news/details/2020-10-27/Petitioner+of+Pornography+Law+Strengthens+Argument>
- Marcetic, B. (2021, Oktober 9). Facebook Harms Its Users Because That's Where Its Profits Are. Jacobin. <https://jacobinmag.com/2021/10/facebook-whistleblower-haugen-profits-addiction-public-utility-regulation-censorship-moderation-zuckerberg>

- Meta. (2018, Mei 23rd). Hard Questions: What's Facebook's Strategy for Stopping False News?. <https://about.fb.com/news/2018/05/hard-questions-false-news/>
- Morozov, E. (2019, Februari 4). Capitalism's New Clothes. *The Baffler*. <https://thebaffler.com/latest/capitalisms-new-clothes-morozov>
- Muhajir, A. (2019). Kenapa Publik Dirugikan Kalau Jurnalis Dijerat UU ITE? Retrieved desember 1, 2021, from <https://www.remotivi.or.id/amatan/516/kenapa-publik-dirugikan-kalau-jurnalis-dijerat-uu-ite>
- Newell, K. (2020). LGBTQ+ Community Leaders in Indonesia: Overcoming Pandemic Hardship. *New Mandala*. <https://www.newmandala.org/lgbtq-community-leaders-in-indonesia-overcoming-pandemic-hardship/> (diakses pada 10 desember 2021)
- Nickel, J. W. (1993). How Human Rights Generate Duties to Protect and Provide. *Human Rights Quarterly*, 15(1).
- Ningtyas, I., Musdalifah, Faisal, E., Maryadi, O., & Afrida, N. (2021). Catatan Akhir Tahun 2021: Kekerasan, Kriminalisasi & Dampak UU Cipta Kerja (Masih Bayangi Jurnalis Indonesia). *Aliansi Jurnalis Independen (AJI)*. https://aji.or.id/upload/article_doc/Catahu_AJI_2021.pdf
- Nugraha, Ricky Mohammad and Laila Afifa (2021). State Uses Virtual Police for Mass Surveillance, SAFENet Says. <https://en.tempo.co/read/1439061/state-uses-virtual-police-for-mass-surveillance-safenet-says> (diakses pada 2 Februari 2022)
- Oates, S. (2020). The easy weaponization of social media: Why profit has trumped security for U.S. companies. *Digital War*, 1(1-3), 117-122. <https://doi.org/10.1057/s42984-020-00012-z>
- Palatino, M. (2015). Will Indonesia's Police Circular on Hate Speech Suppress Freedom of Expression?. <https://advox.globalvoices.org/2015/11/29/will-indonesias-police-circular-on-hate-speech-suppress-freedom-of-expression/#>
- Panjaitan, R. P. (2016, Agustus 18). Mob violence shows Indonesia must act against online hate speech. *The Conversation*. <http://theconversation.com/mob-violence-shows-indonesia-must-act-against-online-hate-speech-63509>
- Pasal 19. (2004). Briefing Note on International and Comparative Defamation Standards.

- Pasal 19. (2017). Germany: The Act to Improve Enforcement of the Law in Social Networks.
- Pasal 19. (2021). Malaysia: Emergency (Essential Powers) (No. 2) Ordinance 2021 (Fake News Ordinance).
- Pasal 19. (2021a). Indonesia: Peraturan Menteri Komunikasi dan Informatika Nomor 5 Tahun 2020 tentang Penyelenggara Sistem Elektronik Swasta (Permenkominfo 5/20).
- Peraturan Daerah Daerah Khusus Ibukota Jakarta Nomor 8 Tahun 2007 tentang Ketertiban Umum.
- Peraturan Daerah Daerah Istimewa Yogyakarta Nomor 2 Tahun 2017 tentang Ketentraman, Ketertiban Umum, dan Perlindungan Masyarakat.
- Peraturan Daerah Kabupaten Buton Nomor 2 Tahun 2020 tentang Penyelenggaraan Ketenteraman, Kedamaian, Ketertiban Umum, dan Perlindungan Masyarakat.
- Price, M. E. "Illegal and Harmful Content", in Fanthome, C. (2002). Global media policy in the new millennium. Indiana University Press.
- Purwati, S. (2018). Unsur Delik Penistaan terhadap Agama dalam Cyber Crime (Analisis Penerapan Pasal 28 ayat (2) UU ITE). Thesis.
- Pusat Studi Agama dan Demokrasi Paramadina. (2018, Agustus 24). Rekayasa Kebencian dalam Kasus Meiliana di Tanjung Balai. tirtoid. <https://tirtoid/rekayasa-kebencian-dalam-kasus-meiliana-di-tanjung-balai-cUEe>
- Putri, D. (2021, Maret 12). Apakah semua ujaran kebencian perlu dipidana? Catatan untuk revisi UU ITE. The Conversation. <http://theconversation.com/apakah-semua-ujaran-kebencian-perlu-dipidana-catatan-untuk-revisi-uu-ite-156132>
- Putusan Pengadilan Negeri Bale Bandung No. 84/Pid.Sus/2021/PN Bbl
- Putusan Pengadilan Negeri Bandung No. 471/Pid.Sus/2020/PN.Bdg
- Putusan Pengadilan Negeri Bantul No. 125/Pid.Sus/2018/PN Btl
- Putusan Pengadilan Negeri Denpasar No. 72/Pid.Sus/2020/PT.Dps
- Putusan Pengadilan Negeri Jakarta Selatan No. 203/Pid.Sus/2019/PN.Jkt.Sel
- Putusan Pengadilan Negeri Jakarta Selatan No. 366/Pid.Sus/2019/PN.JKT.SEL
- Putusan Pengadilan Negeri Jakarta Selatan No. 1333/Pid.Sus/2013/PN.JKT.SEL
- Putusan Pengadilan Negeri Jakarta Timur No. 532/Pid.Sus/2020/PN.Jkt.Tim
- Putusan Pengadilan Negeri Jakarta Utara No. 1537/Pid.B/2016/PN JKT.UTR
- Putusan Pengadilan Negeri Jambi No. 913/Pid.B/2016/PN.Jmb
- Putusan Pengadilan Negeri Jantbo No. 76/Pid.Sus/2021/PN Jth
- Putusan Pengadilan Negeri Jayapura No. 16/Pid.Sus/2020/PN Jap
- Putusan Pengadilan Negeri Kebumen No. 223/Pid.Sus/2018/PN Kbm

- Rosana, Fransisca Christy and Eko Ari Wibowo (2021). 5 Fakta Seputar Terbitnya SKB Pedoman UU ITE, <https://nasional.tempo.co/read/1476302/5-fakta-seputar-terbitnya-skb-pedoman-uu-ite/full&view=ok> (diakses pada 2 Februari 2022).
- Roth, Y & Pickles, N. (2020). Updating our approach to misleading information. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information
- SAFEnet (2020). Pejabat Publik Paling Banyak Laporkan Kasus Pidana UU ITE. <https://databoks.katadata.co.id/datapublish/2021/04/06/safenet-pejabat-publik-paling-banyak-laporkan-kasus-pidana-uu-ite> (diakses pada 2 Februari 2022).
- SAFEnet (2021). Laporan Situasi Hak-hak Digital Indonesia 2020, Represi Digital di Tengah Pandemi.
- Sander, B. (2021). Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law. *European Journal of International Law*, 32(1), 159-194.
- Setianti, L. and Djafar, W. (2017). Tata Kelola Konten Internet di Indonesia: Kebijakan, Praktik, dan Permasalahannya. Policy Brief.
- Setiawan, I. (2021). Who is Responsible for User-Generated Content on Digital Platforms in Indonesia?. Policy Brief (8).
- Sinpeng, A., & Martin, F. R. (2021, Juli 5). Facebook's failure to pay attention to non-English languages is allowing hate speech to flourish. *The Conversation*. <http://theconversation.com/facebooks-failure-to-pay-attention-to-non-english-languages-is-allowing-hate-speech-to-flourish-163723>
- Sinpeng, A., Martin, F. R., Gelber, K., & Shields, K. (2021). Facebook: Regulating Hate Speech in the Asia Pacific. The University of Sydney and The University of Queensland. https://r2pasiapacific.org/files/7099/2021_Facebook_hate_speech_Asia_report.pdf
- Solikhin, A. (2016). Islam, Negara, dan Perlindungan Hak-Hak Islam Minoritas. *Journal of Governance*, 1(1).
- Stevenson, A. (2018, November 6th). Facebook Admits It Was Used to Incite Violence in Myanmar. <https://www.nytimes.com/2018/11/06/technology/myanmar-facebook.html>
- Supreme Court Decision No.183 K/Pid/2010.

- Susanto, E. (2013). Undang-Undang Keterbukaan Informasi Publik dan Penyelenggaraan Pemerintahan. *Komunikator*, 5(1), 54-58.
- Taher, A. P. (2021, September 22). Haris Azhar & Fatia Dipolisikan Luhut, Kuasa Hukum: Salah Alamat. *tirto.id*. <https://tirto.id/haris-azhar-fatia-dipolisikan-luhut-kuasa-hukum-salah-alamat-gjK7>
- Temby, Q. (2019). Disinformation, Violence, and Anti-Chinese Sentiment in Indonesia's 2019 Election. ISEAS Yusof of Ishak Institute Perspective Report, (129). Retrieved from: https://www.iseas.edu.sg/images/pdf/ISEAS_Perspective_2019_67.pdf
- Tempo, "ITE Law's Malleable Terms – The Tale of Suppressive Twins", <https://interaktif.tempo.co/proyek/pasal-karet-uu-ite-sejoli-pembungkam-kritik/index.php>
- The Finery Report. (2021). Journalist from Palopo Sentenced to Prison after Investigation Corruption Case. <https://www.thefineryreport.com/news/2021/11/26/journalist-from-palopo-sentenced-to-prison-after-investigating-corruption-case> (diakses pada 7 Februari 2022).
- Tim Advokasi Bersihkan Indonesia. (2021, September 23). Haris Azhar dan Fatia Maulidiyanti Dilaporkan Luhut Binsar Panjaitan, Ancaman Serius Terhadap Demokrasi dan Kerja-Kerja Pembela Hak Asasi Manusia. *KontraS*. <https://kontras.org/2021/09/23/haris-azhar-dan-fatia-maulidiyanti-dilaporkan-luhut-binsar-panjaitan-ancaman-serius-terhadap-demokrasi-dan-kerja-kerja-pembela-hak-asasi-manusia/>
- Tirto. (2018). "Jerat UU ITE Banyak Dipakai oleh Pejabat Negara", <https://tirto.id/jerat-uu-ite-banyak-dipakai-oleh-pejabat-negara-c7sk> (diakses pada 30 November 2021).
- The Jakarta Post. (2016, Juli 31). False news spread on social media incited riot in N. Sumatra: Police chief. *The Jakarta Post*. <https://www.thejakartapost.com/news/2016/07/31/false-news-spread-on-social-media-incited-riot-in-n-sumatra-police-chief.html>
- Trendacosta, K., & York, J. (2019, Mei 20). TOSsed Out: Highlighting the Effects of Content Rules Online. *Electronic Frontier Foundation*. <https://www.eff.org/deeplinks/2019/05/tossed-out-highlighting-effects-content-rules-online>
- UUD Negara Republik Indonesia Tahun 1945.
- Undang-Undang Nomor 8 Tahun 1999 tentang Perlindungan Konsumen.
- Undang-Undang Nomor 40 Tahun 1999 tentang Pers.

- Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.
- Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.
- Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta.
- United Nations Educational, Scientific and Cultural Organization. (2018). Module 2: Thinking about 'information disorder': formats of misinformation, disinformation, and mal-information. Available at: https://en.unesco.org/sites/default/files/f._jfn_d_handbook_module_2.pdf
- United Nations Developments and Programme & USAID. (2014). Being LGBT In Asia: Indonesia Country Report. <https://www.usaid.gov/documents/2496/being-lgbt-asia-indonesia-country-report-bahasa-language> (diakses pada 10 desember 2021)
- United Nations High Commissioner for Human Rights. (2013). Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred.
- United Nations Strategy and Plan of Action on Hate Speech 2019.
- Valencia, S. (2021, Oktober 28th). Misinformation online is bad in English. But it's far worse in Spanish. <https://www.washingtonpost.com/outlook/2021/10/28/misinformation-spanish-facebook-social-media/>
- Vogelezang, F. (2020, desember 2). Illegal vs Harmful Online Content. <https://www.internetjustsociety.org/illegal-vs-harmful-online-content>
- VOI (2021). After Being Polished, Novel Baswedan was Reported to the KPK. <https://voi.id/en/news/33130/after-being-polished-novel-baswedan-was-reported-to-the-kpk> (diakses pada 2 Februari 2022)
- VOI (2021). Roy Suryo Reveals Mazdjo Pray's Identity, Convicted of Guilty at the Tangerang District Court 2018 Ago. <https://voi.id/en/news/86732/roy-suryo-reveals-mazdjo-prays-identity-was-convicted-of-guilty-at-the-tangerang-district-court-2018-ago> (diakses pada 2 Februari 2022).
- Widhana, D. (2017, Oktober 13). Bagaimana Rasio Dilaporkan Wempy lewat UU ITE. [tirto.id. https://tirto.id/bagaimana-rasio-dilaporkan-wempy-lewat-uu-ite-cyi8](https://tirto.id/bagaimana-rasio-dilaporkan-wempy-lewat-uu-ite-cyi8)
- Wilson, R.A. & Land, M.K. (2021). Hate Speech on Social Media: Content Moderation in Context. *Connecticut Law Review*, 52 (3)

- WM, (2021). "Kemkominfo: 1.971 Hoax COVID-19 Sejak Januari 2020", <https://www.beritasatu.com/nasional/849683/kemkominfo-1971-hoax-covid19-sejak-januari-2020> (diakses pada 2 Februari 2022)
- World Economic Forum. (2021). *Advancing a Digital Safety: A framework to Align Global Action*.
- Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118X.2016.1186713>
- York, J., & Greene, D. (2021, May 25). *Amid Systemic Censorship of Palestinian Voices, Facebook Owes Users Transparency*. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2021/05/amid-systemic-censorship-palestinian-voices-facebook-owes-users-transparency>
- Young, H. (n.d.). *The digital language divide*. The Guardian. Retrieved desember 3, 2021, from <http://labs.theguardian.com/digital-language-divide/>
- Zulli, D., Liu, M., & Gehl, R. (2020). Rethinking the "social" in "social media": Insights into topology, abstraction, and scale on the Mastodon social network. *New Media & Society*, 22(7), 1188–1205. <https://doi.org/10.1177/1461444820912533>



UNIVERSITAS
GADJAH MADA



CENTER FOR DIGITAL SOCIETY



Funded by the
European Union



Center for Digital Society

Faculty of Social and Political Sciences
Universitas Gadjah Mada
Room BC 201-202, BC Building 2nd Floor,
Jalan Socio Yustisia 1
Bulaksumur, Yogyakarta, 55281, Indonesia

Phone : (0274) 563362, Ext. 116
Email : cfds.fisipol@ugm.ac.id
Website : cfds.fisipol.ugm.ac.id



facebook.com/cfdsugm



Center for Digital Society (CfDS)



[cfds_ugm](https://www.instagram.com/cfds_ugm)



[@cfds_ugm](https://wa.me/cfds_ugm)



[@cfds_ugm](https://twitter.com/cfds_ugm)



[CfDS UGM](https://www.youtube.com/CfDSUGM)

