

Kertas Kajian

**Regulasi
Pengaturan
Konten Ilegal
di Media Sosial
di Indonesia**

KERTAS KAJIAN

REGULASI PENGATURAN KONTEN ILEGAL DI MEDIA SOSIAL DI INDONESIA

Penulis : Engelbertus Wendratama
Masduki
Rahayu
Putri Laksmi Nurul Suci

Peninjau : Novi Kurnia
Tim Data Policy and Governance Yayasan Tifa

Perancang sampul & tata letak : gores.pena

Diterbitkan oleh:

Pemantau Regulasi dan Regulator Media (PR2Media)

Jl. Lemponsari Raya, Gg. Masjid RT 9/RW 37 No. 88B

Jongkang Baru, Sariharjo, Sleman, DIY, , 55581

Email: kontak.pr2media@gmail.com

office@pr2media.or.id

Penerbitan kertas kajian ini didukung oleh Yayasan Tifa.

Rekomendasi pengutipan: Wendratama, E., Masduki., Rahayu., & Suci, P. L., (2023).

Kertas Kajian: Regulasi Pengaturan Konten Ilegal di Media Sosial di Indonesia. Yogyakarta: PR2Media.

Publikasi ini tersedia dalam lisensi Creative Commons Attribution-ShareAlike 3.0.

Anda bebas menyalin, mendistribusikan, dan membuat karya turunan dari publikasi ini, asalkan Anda:

- 1) memberikan kredit untuk PR2Media
- 2) tidak menggunakan karya ini untuk tujuan komersial
- 3) mendistribusikan karya apa pun yang berasal dari publikasi ini di bawah lisensi yang identik dengan lisensi ini.

Untuk mengakses teks hukum lengkap dari lisensi ini, silakan kunjungi:

<https://creativecommons.org/licenses/by-sa/3.0/legalcode>

DAFTAR ISI

PENGANTAR	1
A. Tantangan Pengaturan Konten Media Sosial	4
B. Pengertian dan Cakupan Media Sosial	7
C. Pengertian dan Klasifikasi Konten Ilegal	9
D. Tata cara media sosial mengenali dan menentukan konten ilegal.....	16
E. Tata cara media sosial menanggapi aduan konten ilegal.....	16
F. Media sosial menyediakan mekanisme banding	23
G. Dewan Media Sosial	25
H. Laporan tahunan dari media sosial.....	29
I. Audit independen terhadap media sosial	32
PENUTUP	34
REFERENSI	36
LAMPIRAN	41

Kertas kajian ini merupakan konsolidasi aspirasi masyarakat sipil atas perlunya regulasi yang lebih adil dan efektif terkait tata kelola konten media sosial di Indonesia. Para perumus rekomendasi ini merujuk pada tiga sumber. Pertama, hasil riset PR2Media berjudul “Pengaturan Konten Ilegal dan Berbahaya di Media Sosial: Riset Pengalaman Pengguna dan Rekomendasi Kebijakan” (Wendratama et al., 2023). Kedua, penyelenggaraan diskusi kelompok terarah dengan para pihak, terutama pakar hukum, praktisi media, dan kelompok masyarakat sipil di Jakarta, 6–7 Juni 2023. Ketiga, penelusuran referensi terkait pengaturan tata kelola media di Indonesia dan beberapa negara lain, misalnya, kajian yang telah dilakukan oleh Center for Digital Society, FISIPOL Universitas Gadjah Mada berjudul “Regulating Harmful Content in Indonesia: Legal Frameworks, Trends, and Concerns” (2022) dan oleh Article 19 berjudul “Content Moderation and Local Stakeholders in Indonesia” (2022).

Kertas kajian ini merupakan bagian integral dari advokasi kebijakan untuk menciptakan tata kelola media sosial yang lebih sehat di Indonesia. Format rekomendasi ini dibuat sesederhana mungkin, substantif, dan dilengkapi usulan pasal-pasal untuk undang-undang, agar lebih mudah dicermati oleh pembuat kebijakan.

Mencermati struktur regulasi setingkat undang-undang di Indonesia hingga 2023, dan memperhatikan kondisi empiris ekonomi politik penyusunan kebijakan, maka tim perumus dan organisasi masyarakat sipil yang terlibat dalam kerja advokasi ini menargetkan dua luaran kebijakan sebagai hasil dari kertas kajian ini.

Pertama, dalam jangka pendek, kertas kajian ini diajukan dalam pembahasan revisi Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), khususnya revisi Pasal 15 tentang tanggung jawab penyelenggara sistem elektronik, yang hingga Agustus 2023 masih dibahas oleh Komisi I DPR RI dan pemerintah. Usulan ini telah disampaikan PR2Media dalam Rapat Dengar Pendapat Umum (RDPU) yang diadakan Panja Revisi UU ITE Komisi I DPR RI pada 23 Agustus 2023 (PR2Media, 2023).

Kedua, dalam jangka panjang, dokumen ini diajukan sebagai bagian dari reformasi kebijakan tata kelola internet di Indonesia, yang bertolak dari revisi total UU ITE menuju undang-undang yang lebih komprehensif dan adil dalam mengatur tata cara berbagai platform digital mengatasi konten ilegal.

Apa latar belakang pengajuan dokumen ini? Merujuk hasil riset PR2Media pada 2023, di Indonesia, cara perusahaan media sosial menangani konten ilegal untuk pengguna di Indonesia masih kurang efektif, transparan, dan berperspektif hak asasi manusia, dibandingkan yang mereka lakukan di negara-negara demokrasi maju. Sebagai perbandingan, Uni Eropa telah memiliki Digital Services Act (mulai berlaku efektif awal 2024) yang menerapkan standar tinggi bagi transparansi dan akuntabilitas berbagai platform digital, terutama platform dengan jumlah pengguna besar, dalam menangani konten ilegal.

Saat ini, terdapat sejumlah regulasi yang mengatur konten internet di Indonesia: UU ITE No. 19/2016, Peraturan Pemerintah No. 71/2019 tentang Penyelenggaraan Sistem Transaksi Elektronik, dan Permenkominfo No. 5 Tahun 2020 tentang Penyelenggaraan Sistem Elektronik Lingkup Privat. Namun, seperangkat regulasi ini masih belum memadai untuk mendorong penguatan peran pemerintah selaku regulator, serta transparansi dan akuntabilitas platform digital dalam penanganan konten ilegal di satu sisi, dan menjaga hak publik atas kebebasan berekspresi di internet di sisi lain.

Selain itu, platform media sosial masih sangat bergantung pada *artificial intelligence* (AI) dalam memutuskan sebuah konten atau akun perlu dihapus atau tidak. Jumlah moderator manusia yang dipekerjakan platform untuk memoderasi konten masih minim, demikian pula transparansi proses moderasi konten. Masa kampanye untuk Pemilihan Umum (Pemilu) 2024 berlangsung sejak November 2023, yang berpotensi meningkatkan penyebaran disinformasi dan ujaran kebencian yang memperburuk proses demokrasi. Berkaca dari kekacauan informasi yang terjadi pada proses politik elektoral pada tahun-tahun sebelumnya dan dampaknya pada proses demokrasi dan kohesi sosial, terobosan regulasi yang secara detail mengatur standar transparansi dan akuntabilitas moderasi konten ilegal oleh platform media sosial menjadi semakin urgent.

Konten ilegal adalah segala jenis konten yang menurut peraturan perundang-undangan di Indonesia terlarang, misalnya pornografi (melanggar UU No. 44/2008), perjudian (KUHP, UU ITE No. 19/2016), ujaran kebencian (KUHP, UU No. 19/2016, UU No. 40/2008, Surat Edaran Kapolri SE/6/X/2015), dan pencemaran nama baik serta kabar bohong (KUHP, UU No. 19/2016). Harus diakui, definisi konten ilegal di Indonesia

masih abu-abu karena sifatnya sangat kontekstual, sehingga diperlukan regulasi yang memandu moderasinya oleh platform.

Berbekal metodologi yang kuat, temuan yang dihasilkan riset PR2Media ini memiliki legitimasi untuk memberikan rekomendasi terhadap revisi UU ITE dan penyusunan peraturan perundang-undangan baru yang mengatur tanggung jawab penyelenggaraan media sosial di Indonesia. Sebagai contoh, survei PR2Media pada awal 2023 itu menunjukkan bahwa banyak pengguna media sosial di Indonesia yang tidak puas dengan cara perusahaan media sosial memoderasi konten dan melayani aduan mereka. Temuan ini bisa menjadi legitimasi bagi pemerintah dan Dewan Perwakilan Rakyat untuk menemukan langkah yang lebih baik dalam pengaturan konten dan layanan aduan pengguna selaku warga negara. Riset tersebut juga diperkuat dengan perspektif para pemangku kepentingan dalam tahap wawancara, yang memperkaya argumentasi perlunya terobosan kebijakan dalam pengaturan konten media sosial di Indonesia.

Penyusunan kajian ini telah melihat (harmonisasi) regulasi sejenis yang sudah ada, yaitu UU ITE No. 19/2016, PP No. 71/2019, dan Permenkominfo No. 5 Tahun 2020. Sehingga, usulan regulasi ini akan melengkapi dan memperkuat regulasi yang sudah ada. Riset PR2Media menunjukkan revisi UU ITE secara total adalah jalur mewujudkan payung hukum yang lebih kuat dan komprehensif bagi pengaturan konten ilegal.

Di samping itu, tim peneliti juga menilai revisi undang-undang ini perlu dilihat dalam kerangka tata kelola digital yang lebih luas. Misalnya, saat ini ada wacana pengaturan konten *video on demand* (seperti Netflix) dan *video sharing platform* (seperti YouTube) dalam pembahasan revisi UU Penyiaran. Uni Eropa mengaturnya melalui Audio-visual Media Services Act. Gagasan ini selaras dengan perlunya pengaturan aspek lain dalam jagat digital seperti persaingan usaha antar platform digital (Digital Markets Act di Uni Eropa) dan kecerdasan buatan (AI Act di Uni Eropa).

Dokumen ini terdiri dari tiga aspek kajian. Pertama, penjelasan mendasar terkait tantangan pengaturan konten ilegal di media sosial, definisi dan cakupan media sosial, definisi dan klasifikasi konten ilegal. Kedua, tata cara media sosial mengenali dan menentukan konten ilegal, tata cara dalam menanggapi aduan konten ilegal, dan perlunya penyediaan mekanisme banding oleh pengelola platform. Ketiga, inisiatif pembentukan Dewan Media Sosial (DMS), kewajiban bagi pengelola media sosial untuk menerbitkan laporan transparansi tahunan mengenai keputusan moderasi konten, dan audit independen terhadap laporan ini. Sebagai bagian dari upaya memberikan rekomendasi kebijakan yang lebih konkret kepada pemerintah dan DPR, kajian ini melampirkan usulan perubahan terhadap pasal 15 pada UU ITE mengenai tanggung jawab penyelenggara sistem elektronik di Indonesia.

A. Tantangan Pengaturan Konten Media Sosial

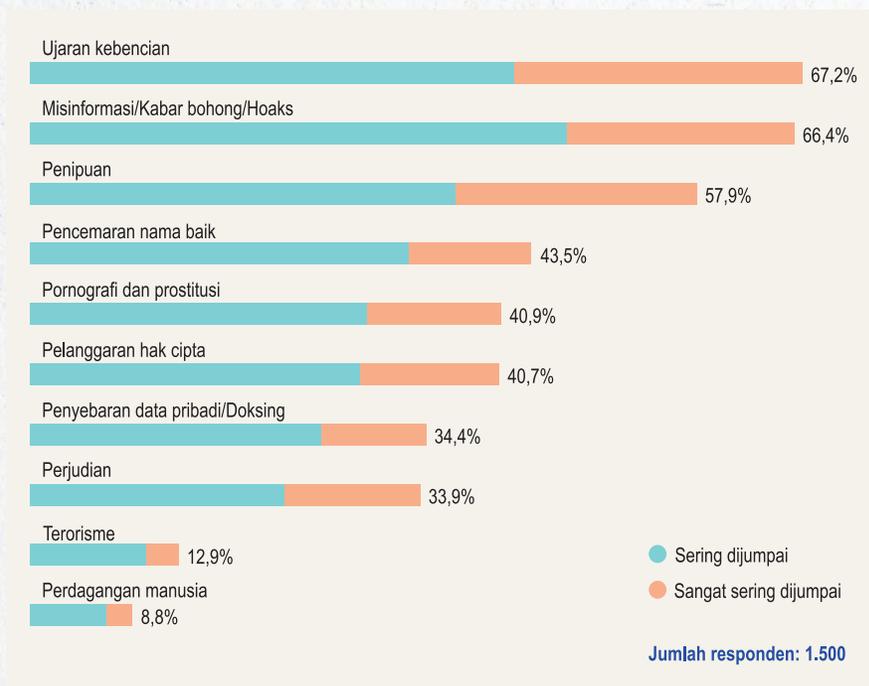
1. Kondisi penyebaran konten ilegal di media sosial saat ini

Penyebaran konten ilegal melalui media sosial yang dialami oleh masyarakat Indonesia telah lama menjadi kepedulian berbagai pihak. Berdasarkan data Direktorat Jenderal Aplikasi Informatika Kementerian Komunikasi dan Informatika hingga 6 Maret 2023, konten ilegal yang dilaporkan Kementerian Komunikasi dan Informatika kepada penyelenggara sistem elektronik mencapai hampir 1,4 juta konten, terbanyak ada di Twitter, diikuti Meta (Facebook dan Instagram), dengan jenis terbanyak adalah pornografi, perjudian, dan penipuan (Kementerian Komunikasi dan Informatika, 2023).

Sementara itu, riset PR2Media pada awal 2023 berjudul “Pengaturan Konten Ilegal dan Berbahaya di Media Sosial: Riset Pengalaman Pengguna dan Rekomendasi Kebijakan” (Wendratama et al., 2023) juga menunjukkan kondisi yang memprihatinkan. Seluruh responden riset itu (1.500 pengguna media sosial di 38 provinsi Indonesia) menyatakan sering menjumpai konten ilegal di media sosial, dengan tiga jenis yang paling sering dijumpai adalah ujaran kebencian, misinformasi/kabar bohong/hoaks, dan penipuan.

Urutan konten ilegal yang sering dijumpai masyarakat Indonesia di media sosial tampak dalam gambar berikut ini.

Gambar 1. Jenis Konten Ilegal yang Paling Sering Dijumpai Masyarakat Indonesia



Perlu dicatat bahwa beberapa jenis konten ilegal tersebut diartikan sebagai konten berbahaya (tidak boleh dikriminalisasi) oleh penyelenggara media sosial dan berbagai regulasi di negara lain, termasuk sejumlah instrumen PBB, yaitu misinformasi, (sebagian) ujaran kebencian, dan pornografi dewasa (bukan anak). Meski demikian, riset PR2Media tersebut melakukan pengelompokan di atas supaya sesuai dengan peraturan perundang-undangan yang berlaku di Indonesia.

Seiring makin banyaknya pengguna media sosial di Indonesia, dengan variasi konten yang makin beragam pula, risiko penyebaran konten ilegal pun semakin tinggi. Namun, sebagaimana ditunjukkan oleh riset PR2Media yang juga mewawancarai beragam pemangku kepentingan, instrumen regulasi yang ada saat ini dinilai belum memadai untuk mengatasi tantangan tersebut. Tantangan regulasi tersebut akan diuraikan di bawah ini.

2. Tantangan regulasi di Indonesia

Di Indonesia, pengaturan atas moderasi konten ilegal di media sosial baru mengatur wewenang yang dimiliki pemerintah dalam mencegah penyebarluasan dan penggunaan konten ilegal dan tata caranya, tapi belum mengatur tata cara penyelenggara media sosial mengatur konten ilegal, padahal penyelenggara media sosial memiliki sumber daya yang lebih besar dan sesuai untuk itu.

Regulasi yang mengatur konten media sosial (dan internet secara keseluruhan) adalah UU No. 19 Tahun 2016 tentang Perubahan atas UU No. 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE), yang diturunkan menjadi Peraturan Pemerintah No. 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik (PP No. 71/2019), dan Peraturan Menteri Komunikasi dan Informatika No. 5 Tahun 2020 tentang Penyelenggaraan Sistem Elektronik Lingkup Privat (Permenkominfo No. 5/2020).

Ada dua tantangan utama yang dihadapi UU ITE dalam mengatur penyebaran konten ilegal di media sosial, yaitu:

1. Definisi konten ilegal belum detail sehingga sering muncul perbedaan penafsiran antara pemerintah, penyelenggara media sosial, dan masyarakat tentang konten yang dilarang, misalnya soal definisi “mengganggu ketertiban umum”.
2. Pengaturan kewajiban dan sanksi bagi penyelenggara media sosial dalam menangani konten ilegal belum detail. Termasuk di sini adalah semua penyelenggara sistem elektronik (mesin pencari, layanan penyimpanan digital, lokapasar (*marketplace*), aplikasi percakapan, dll) disamaratakan oleh UU ITE, padahal karakteristik dan jumlah penggunaannya (pengaruhnya di masyarakat) berbeda-beda.

Dua hal itu menjadi masalah besar mengingat yang berkuasa dalam mengatur (misalnya: menghapus) konten di media sosial adalah penyelenggara media sosial yang bersangkutan, karena pemerintah dan pengguna hanya bisa menyampaikan aduan kepada mereka.

Pengaturan (moderasi) konten adalah kegiatan penyelenggara media sosial dalam menilai, menandai, membatasi akses, menurunkan popularitas, dan/atau menghapus konten ilegal dan berbahaya atau akun yang mengunggah konten tersebut untuk melindungi penggunanya. Dalam melakukan penilaian atas konten, platform media sosial memanfaatkan moderator manusia dan/atau kecerdasan buatan (*artificial intelligence/AI*).

Jika penyelenggara media sosial tidak mengabulkan aduan dari pemerintah, hanya ada satu pilihan ekstrem yang bisa diambil pemerintah, yaitu memblokir akses pengguna di Indonesia ke platform tersebut secara total. Langkah ini kontraproduktif dan otoriter karena masih banyak konten yang legal dan bermanfaat bagi warga di platform tersebut.

Karena kesulitan tersebut, dengan menyadari besarnya kuasa penyelenggara media sosial terhadap pengaturan (moderasi) konten dan dampaknya bagi masyarakat, mulai muncul regulasi yang mengatur kewajiban penyelenggara media sosial terhadap konten ilegal di platformnya.

Uni Eropa, misalnya, telah memiliki Digital Services Act (disahkan pada 2022, diterapkan pada 2024), yang mengatur cara seluruh jenis penyelenggara layanan internet mengatasi konten ilegal di platformnya. Platform yang diatur mulai dari mesin pencari, layanan jual beli, hingga media sosial terbuka. Kewajiban mereka pun berbeda, disesuaikan dengan jumlah penggunanya di Uni Eropa, dengan platform besar memiliki kewajiban yang lebih banyak pula. Khusus untuk media sosial terbuka berukuran besar¹, kewajiban mereka juga lebih banyak, termasuk di dalamnya adalah membuat laporan tentang cara platform media sosial melakukan pengaturan (moderasi) konten dan memitigasi risiko yang muncul akibat desain dan penggunaan platform (European Commission, 2023).

¹ Platform dengan pengguna aktif bulanan lebih dari 10% populasi Uni Eropa (45 juta). Pada April 2023, mereka menentukan platform media sosial terbuka ini adalah YouTube, Facebook, TikTok, Instagram, Twitter, dan LinkedIn (European Commission, 2023).

Regulasi di Uni Eropa tersebut mengatur cara platform memoderasi konten ilegal, sehingga pemerintah tidak langsung mengatur kontennya. Cara ini dinilai lebih efektif karena penyelenggara media sosial memiliki sumber daya yang lebih lengkap dan sesuai untuk itu. Pemerintah tinggal mengawasi kepatuhan penyelenggara media sosial terhadap regulasi itu.

Seperti regulasi yang sudah ada di Jerman dan Prancis sebelumnya, Digital Services Act (DSA) mengatur cara penyelenggara mengatur konten, sehingga negara atau pemerintah tidak langsung mengatur kontennya. Cara ini dinilai lebih efektif karena penyelenggara media sosial memiliki sumber daya yang lebih lengkap dan sesuai untuk itu. Dengan begitu, pemerintah tinggal mengatur dan mengawasi praktik yang dilakukan oleh penyelenggara media sosial—sesuatu yang juga diharapkan oleh Direktur Jenderal Aplikasi dan Informatika, Samuel Abrijani Pangerapan, dalam wawancara dengan peneliti PR2Media pada Maret 2023 (Wendratama et al., 2023). Negara lain yang sedang membahas undang-undang seperti itu adalah Inggris dan Singapura, dalam bentuk Online Safety Bill.

B. Pengertian dan Cakupan Media Sosial

Pengertian

Definisi media sosial cukup beragam dan tidak ada kesepakatan tunggal di antara para akademisi dan praktisi. Meski demikian, berdasarkan riset Aichner et al. (2021) yang memetakan sejarah definisi media sosial sejak 1994 hingga 2019 di berbagai publikasi akademik, penulis kajian ini mendefinisikan media sosial sebagai “sistem elektronik berbasis internet yang memungkinkan para penggunanya saling mengirimkan konten dalam jaringan yang bersifat terbuka dan diatur oleh penyelenggara media sosial”.

Di sini penekanannya adalah “para pengguna bisa saling mengirimkan konten” dalam jaringan “bersifat terbuka”, seperti YouTube, Facebook, TikTok, Instagram, dan LinkedIn, bukan “bersifat tertutup” yang dikripsi seperti WhatsApp, Telegram, dan Signal.

Tentu, penyebaran konten ilegal menurut hukum Indonesia juga bisa terjadi pada platform selain media sosial, seperti mesin pencari, layanan penyimpanan digital, lokapasar (*marketplace*), dan aplikasi percakapan. Namun, karena riset yang dilakukan PR2Media (2023) baru terbatas pada media sosial terbuka, maka usulan regulasi ini pun berfokus pada media sosial terbuka. Idealnya, regulasi yang mengatur kewajiban

penyelenggara sistem elektronik terkait penyebaran konten ilegal juga mencakup jenis apa pun (seperti yang dilakukan oleh UU ITE), tapi kewajiban tiap platform perlu disesuaikan dengan karakteristiknya masing-masing, seperti yang dilakukan oleh Digital Services Act di Uni Eropa.

Cakupan

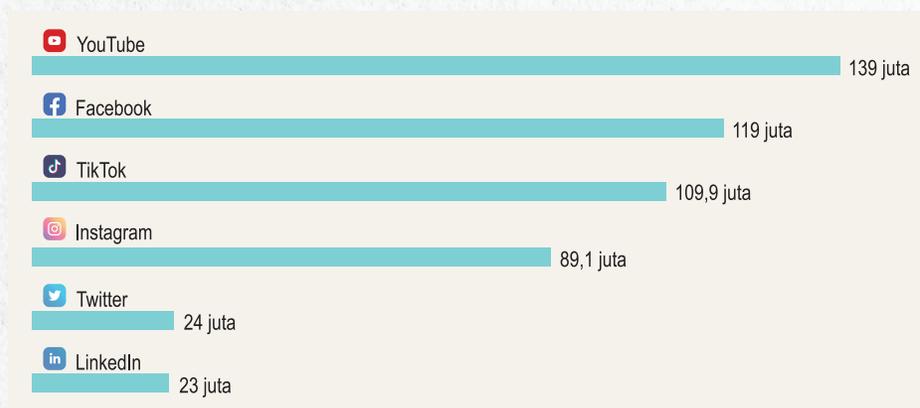
Regulasi yang kami usulkan dalam pengaturan konten media sosial ini berfokus pada media sosial yang memiliki jumlah pengguna banyak, dengan pertimbangan bahwa mereka memiliki dampak yang relatif besar terhadap masyarakat Indonesia. Pertimbangan jangkauan seperti ini juga diadopsi oleh Digital Services Act, dengan platform yang memiliki pengguna besar akan memiliki kewajiban yang lebih banyak pula. Kelompok platform yang memiliki daftar kewajiban paling banyak banyak dan detail adalah yang memiliki pengguna aktif bulanan lebih dari 10% populasi Uni Eropa (45 juta pengguna). Pada April 2023, Komisi Eropa mengumumkan platform media sosial terbuka yang termasuk kelompok ini adalah YouTube, Facebook, TikTok, Instagram, Twitter, dan LinkedIn (European Commission, 2023).

PR2Media menilai, angka 10% populasi untuk konteks Indonesia merupakan jumlah yang terlalu besar, yaitu 27 juta pengguna (jumlah penduduk Indonesia berdasarkan data BPS tahun 2023 sebesar 278 juta jiwa). Padahal, sebagai contoh, jumlah pengguna Twitter di Indonesia pada Januari 2023 sebanyak 24 juta orang (Kemp, 2023).

Oleh karena itu, menurut kami, batas yang rasional adalah 20 juta, sehingga usulan regulasi ini hanya berkenaan dengan media sosial yang memiliki jumlah pengguna di Indonesia minimal 20 juta orang. Tentunya batas itu selalu bisa diubah sesuai dinamika di masyarakat, sebagaimana Komisi Eropa yang akan selalu memperbarui data platform yang dianggap sebagai *very large online platform* (VLOP), yang memiliki kewajiban paling banyak dalam DSA.

Dengan ambang batas 20 juta pengguna ini, maka media sosial yang dimaksud dalam usulan regulasi ini adalah: YouTube, Facebook, TikTok, Instagram, Twitter, dan LinkedIn.

Gambar 2. Jumlah pengguna media sosial di Indonesia pada awal 2023



C. Pengertian dan Klasifikasi Konten Ilegal

Tantangan berikutnya adalah definisi konten ilegal di hukum Indonesia yang sangat luas, yang juga mencakup “konten berbahaya” menurut pengertian penyelenggara media sosial. Konten berbahaya adalah konten yang legal, tapi dapat melukai atau merugikan pengguna secara fisik maupun psikologis.

Sebagai contoh, menurut definisi penyelenggara media sosial dan sejumlah instrumen PBB (United Nations Human Rights Office of the High Commissioner, 2023), hoaks, ujaran kebencian, pornografi orang dewasa, dan perjudian bukanlah konten ilegal, melainkan sebagai konten berbahaya yang tidak boleh dikriminalisasi. Namun, menurut hukum Indonesia, semua itu merupakan konten ilegal. Dengan kata lain, publikasi dan penyebaran konten berbahaya merupakan tindak pidana menurut regulasi Indonesia, sehingga semua konten berbahaya dapat dituntut secara pidana. Karena itu, hanya ada dua mekanisme yang berlaku dalam menangani konten ilegal maupun berbahaya di bawah hukum Indonesia: mekanisme pengadilan (penuntutan pidana, dan untuk beberapa kasus, tuntutan perdata) dan mekanisme non-pengadilan (penyelesaian sengketa dan tindakan administratif lainnya) (Rahman et al., 2022).

Akibatnya, seperti terjadi selama ini, beragam pasal dalam UU ITE sering dipakai untuk mengkriminalisasi warga yang menyampaikan kritik atau pendapat, sehingga mengekang kebebasan berekspresi warga negara, yang sesungguhnya dilindungi oleh UUD 1945.

Selain memberikan ancaman pidana terhadap berbagai jenis konten berbahaya di media sosial (yang spektrumnya sangat luas), UU ITE juga menimbulkan multitafsir dan kontroversi akibat beberapa istilah yang memiliki arti terlalu luas, seperti moralitas dan ketertiban umum.

Sejumlah pelaksanaan peraturan yang berhubungan dengan konten ilegal pun dibuat untuk “memandu” penegak hukum, seperti Surat Keputusan Bersama Pedoman Kriteria Implementasi UU Informasi dan Elektronika (UU ITE) oleh Menkominfo, Jaksa Agung, dan Kapolri (Keputusan Bersama, 2021) dan Surat Edaran Kepala Kepolisian Negara Republik Indonesia Nomor SE/6/X/2015 tentang Penanganan Ujaran Kebencian. Meski demikian, berbagai pihak menilai pendefinisian konten yang melanggar peraturan perundang-undangan masih problematis.

Sebagai contoh, Surat Edaran Kapolri tahun 2015 itu mengategorikan penghinaan, pencemaran nama baik, dan penistaan sebagai bagian dari ujaran kebencian serta tidak mendefinisikan secara jelas karakteristik tindakan-tindakan tersebut (Palatino, 2015). Sebagai perbandingan, jika mengacu pada sejumlah kesepakatan internasional (Gagliardone et al., 2015), ujaran kebencian terbatas pada “serangan terhadap seseorang atau kelompok orang berdasarkan ‘karakteristik inheren’ yang dimilikinya seperti suku, agama, ras, nasionalitas, dan gender”.

1. Pertimbangan hukum untuk ujaran kebencian, hoaks, dan pencemaran nama baik

Salah satu jenis konten ilegal yang paling problematik di Indonesia adalah ujaran kebencian. Survei terhadap 1.500 pengguna media media sosial dari 38 provinsi di Indonesia yang dilakukan PR2Media dalam riset berjudul “Pengaturan Konten Ilegal dan Berbahaya di Media Sosial” (Wendratama et al., 2023) menunjukkan, jenis konten ilegal dan berbahaya yang paling sering dijumpai oleh responden adalah ujaran kebencian. Persepsi para responden ini tidak bisa dilepaskan dari fakta hukum Indonesia yang mendefinisikan ujaran kebencian secara sangat luas, sehingga mereka pun menemui banyaknya kasus hukum terkait ujaran kebencian, dengan spektrum kasus yang sangat luas dan mengancam hak-hak berekspresi warga.

Karena itu, kajian ini mengusulkan supaya regulasi di Indonesia mengadopsi tiga klasifikasi ujaran kebencian. Rujukannya adalah dokumen Rabat Plan of Action yang diadopsi oleh Kantor Komisararis Tinggi PBB untuk Hak Asasi Manusia (United Nations High Commissioner for Human Rights, 2013) yang mengelompokkan ujaran kebencian menjadi tiga, yaitu:

1. Penyampaian pendapat yang harus diancam pidana.

Penyampaian pendapat yang harus diancam pidana adalah (1) menghasut perbuatan diskriminasi, permusuhan, atau kekerasan terhadap suatu bangsa,

ras, agama dan (2) membenarkan atau menganjurkan kebencian terhadap ras dan diskriminasi dalam bentuk apapun. Di sini pelaku bermaksud dan sengaja untuk menggerakkan orang lain. Dua poin itu diatur oleh dua konvensi PBB, yang sudah diratifikasi oleh Indonesia, yaitu International Covenant on Civil and Political Rights (International Covenant on Civil and Political Rights, 1966) dan International Convention on the Elimination of All Forms of Racial Discrimination (International Convention on the Elimination of All Forms of Racial Discrimination, 1965).

2. Penyampaian pendapat yang dapat diancam dengan sanksi administrasi atau digugat secara perdata.

Ujaran kebencian yang termasuk kategori ini adalah ucapan yang mengandung kebencian didasarkan pada Pasal 19 Ayat 3 International Covenant on Civil and Political Rights (ICCPR) yang mengatur bahwa hak mengeluarkan pendapat dapat dibatasi untuk melindungi hak dan reputasi orang lain, keamanan negara atau ketertiban umum, kesehatan publik, atau untuk kepentingan moral.

3. Penyampaian pendapat yang tidak dapat diancam sanksi apa pun.

Pendapat yang tidak dapat diancam sanksi apapun adalah semata sifat intoleransi dan perasaan tidak suka yang dimiliki seseorang. Untuk mengatasinya, jalan terbaik adalah melalui edukasi. Salah satu bentuk edukasi yang gencar dilakukan adalah kampanye literasi digital oleh pemerintah dan berbagai pihak.

Dengan demikian, tidak semua ujaran kebencian bisa dipidanakan (Putri, 2021) sehingga tidak berpotensi merugikan hak berekspresi yang merupakan hak asasi warga negara. Jika kita melihat kasus-kasus hukum ujaran kebencian yang ada di Indonesia, mayoritas kasus sebenarnya masuk ke kategori nomor 2 dan 3, sehingga tidak perlu dipidana.

Upaya berikutnya untuk bisa mendorong kesamaan tafsir dan meminimalkan pelanggaran hak warga adalah adopsi uji ambang batas, yang juga berdasarkan Rabat Plan of Action (United Nations High Commissioner for Human Rights, 2013). Ambang batas (*threshold*) ini sejatinya diperuntukkan bagi konten ujaran kebencian, tapi juga bisa diadopsi dalam penentuan apakah suatu konten apa pun layak dipidana atau tidak, terutama terkait tuduhan hoaks dan pencemaran nama baik. Enam aspek itu diuraikan sebagai berikut.

UJI ENAM AMBANG BATAS

1. Konteks

Suatu konten perlu dianalisis dalam konteks sosial dan politik ketika konten itu dibuat dan dibagikan.

2. Pembicara

Posisi atau status pembicara atau pelaku di masyarakat perlu dipertimbangkan, terutama posisi orang atau organisasi itu di hadapan audiens yang menjadi sasaran konten itu.

3. Niat

Kelalaian dan kecerobohan tidak cukup untuk membuat tindakannya dipidanakan. Harus ada unsur hasutan dan provokasi (yang menunjukkan niat) dari orang itu terhadap audiensnya, bukan semata tindakan penyebarluasan konten.

4. Isi dan bentuk

Analisis terhadap isi dan bentuk pesan akan menilai kadar provokasi atau hasutan tersebut, termasuk pesan yang disampaikan bersifat langsung (eksplisit, lugas) atau tidak, dan gaya penyampaianya seperti apa (misalnya konten satir).

5. Jangkauan

Hal ini mengacu pada jangkauan audiens, yakni besar atau kecilnya audiens. Semakin besar audiens, maka semakin besar pula potensi risikonya bagi publik.

6. Kemungkinan risiko

Ini mengacu pada kemungkinan yang masuk akal bahwa pesan tersebut akan berhasil menghasut adanya tindakan terhadap kelompok yang disasar, dengan sebab-akibat itu bersifat langsung.

PR2Media menilai, enam ambang atas ini juga bisa dipakai dalam penentuan konten ilegal selain ujaran kebencian, terutama lima ambang batas pertama. Sebagai contoh, terkait penyebaran hoaks dan pencemaran nama baik, yang merupakan tindakan ilegal di Indonesia, lima ambang batas pertama juga bisa diterapkan untuk mengurangi multitafsir dan kasus pemidanaan.

Dengan mengacu pada dua model tingkatan di atas (klasifikasi dan ambang batas), PR2Media mengusulkan supaya bahwa konten ilegal yang dianggap sebagai konten berbahaya (terutama hoaks dan pencemaran nama baik) tidak harus selalu dipidanakan, tapi mempertimbangkan juga intensitas dan jangkauannya.

Tanggung jawab secara hukum atau liabilitas penyelenggara media sosial terkait konten yang diunggah pengguna menjadi perdebatan di banyak negara. Di Amerika

Serikat misalnya, muncul wacana merevisi Section 230 dalam Communications Decency Act (1996), yang memberikan imunitas kepada platform media sosial dari liabilitas hukum sehingga mereka tidak bisa dituntut atas konten yang diunggah pengguna. Menurut undang-undang itu, platform media sosial tidak bisa diberlakukan sama dengan penerbit berita, yang bisa dituntut atas konten yang mereka terbitkan. Meski demikian, ada konten pengguna yang karenanya platform bisa dituntut secara hukum, yaitu pelanggaran terhadap hak cipta, hukum kriminal Federal, dan perdagangan seks anak atau secara paksa (US Congress, 2019). Menanggapi rencana revisi itu, Mark Zuckerberg menyatakan persetujuannya bahwa perlu ada liabilitas hukum untuk penyelenggara media sosial terkait “sebagian” konten unggahan pengguna, tapi menolak jika penyelenggara media sosial harus bertanggung jawab secara hukum untuk seluruh konten yang diunggah pengguna. Hingga saat ini, proses revisi masih berlanjut, dengan Presiden Joe Biden mendorong Kongres untuk segera menyelesaikan revisi itu (Morrison, 2023).

Secara umum, rezim liabilitas platform yang berlaku di berbagai negara dapat dikelompokkan menjadi empat, yaitu (1) model liabilitas ketat (*strict liability*), di mana platform bertanggung jawab penuh atas konten pihak ketiga; (2) model *knowledge-based liability*, di mana platform dibebaskan dari sanksi kecuali ia mengetahui adanya konten ilegal di layanannya dan gagal untuk menindaknya; (3) model *fault-based liability*, di mana platform dapat dikenakan sanksi jika dianggap gagal melakukan pencegahan atas penyebaran konten yang melanggar; dan (4) model imunitas luas (*broad immunity from liability*) (Frosio, 2021).

Kajian ini menilai, model liabilitas yang berbeda perlu diberlakukan atas konten ilegal dan berbahaya. Misalnya, model *knowledge-based liability* lebih tepat digunakan untuk konten ilegal mendesak (kecuali misinformasi dan ujaran kebencian). Hal ini sama dengan yang diadopsi oleh regulasi Digital Services Act. Definisi *knowledge* atau pengetahuan yang dimaksud hanya terbatas pada laporan dari pelapor terpercaya (*trusted flagger*) dan keputusan pengadilan (supaya penyelenggara media sosial tidak harus melakukan pemantauan umum, menyaring, dan/atau menghindari penyalahgunaan mekanisme pelaporan).

Keberadaan pelapor terpercaya (*trusted flagger*), yaitu organisasi yang memiliki kompetensi dalam mengenali dan menandai konten ilegal, memiliki posisi penting dalam ekosistem moderasi konten. Laporan dari pelapor terpercaya ini perlu mendapat prioritas dari penyelenggara media sosial karena mereka sudah terlatih membuat laporan yang lengkap dan terpercaya. Organisasi ini mendapat pelatihan dari penyelenggara media sosial dan menjadi mitra mereka. Salah satu contohnya adalah ECPAT Indonesia (organisasi yang berfokus pada penghapusan kekerasan seksual terhadap anak) yang pernah menjadi pelapor terpercaya bagi Twitter.

2. Pertimbangan hukum untuk pencemaran nama baik dan penghinaan

Kasus pencemaran nama baik dan penghinaan lekat dengan urusan privat. Karena itu, pengaturannya pun sebaiknya berbeda, yaitu melalui pendekatan *notice-to-notice*, di mana penyelenggara media sosial dapat memberikan *counter-notice* jika mereka menerima aduan dari pengguna, lalu keduanya memutuskan akan lanjut ke proses pengadilan atau tidak.

Salah satu perubahan dalam revisi Undang-Undang Informasi dan Transaksi Elektronik pada 2016 adalah pada Pasal 27, yaitu untuk menghindari multitafsir, maka tafsir pencemaran nama baik dikembalikan ke KUHP (lama) Pasal 310 dan 311, yaitu:

1. Informasi tersebut harus bertujuan menyerang kehormatan atau nama baik seseorang dan ini tersiar secara umum (diketahui orang banyak).
2. Perbuatan tidak dilakukan demi kepentingan umum atau bukan untuk membela diri (terpaksa).

Dengan kata lain:

1. Jika yang dituduhkan tersebut benar tetapi tersiarnya tuduhan tersebut bukan demi kepentingan umum atau membela diri sendiri (Pasal 310 Ayat 3 KUHP), maka dapat dipidana berdasarkan Pasal 310 Ayat 1 KUHP tentang penistaan atau penghinaan.
2. Jika tuduhan perbuatan tersebut terbukti tidak benar, maka dapat dipidana dengan Pasal 311 Ayat 1 KUHP mengenai fitnah.

Dengan demikian, meskipun suatu pernyataan (konten) itu benar atau sesuai fakta, orang yang menyampaikan pernyataan tersebut di platform elektronik masih dapat dipidanakan dengan pasal penistaan atau penghinaan.

Pencemaran nama baik dan penghinaan sangat erat dengan konsep privasi. Dalam konteks penyebaran informasi, privasi mengacu pada hak seseorang untuk merahasiakan kehidupan dan informasi pribadinya atau membuatnya hanya diketahui sedikit orang saja (Wendratama, 2021). Sayangnya, di Indonesia belum ada diskusi tentang perbedaan

hak privasi antara warga biasa dan tokoh publik (*public figure*), sebagaimana di negara-negara lain. Preseden hukum Amerika Serikat, misalnya, membedakan privasi yang dimiliki warga negara biasa dan privasi yang dimiliki oleh tokoh publik. Di AS, tingkat privasi figur publik lebih rendah karena mereka bekerja untuk melayani publik atau mendapat uang dari perhatian publik. Empat aspek dalam preseden hukum AS (Yanisky-Ravid & Lahav, 2006) untuk memenuhi definisi tokoh publik tersebut adalah:

1. Akses dan kontrol terhadap media (termasuk di sini adalah politikus dan siapa pun yang sering diliput media).
2. Memiliki peranan dalam kehidupan publik, yaitu orang yang memegang jabatan publik dan/atau menggunakan sumber daya publik, termasuk semua orang yang berperan dalam kehidupan publik, baik dalam politik, ekonomi, seni, bidang sosial, olahraga, atau bidang apa pun.
3. Siapa pun yang secara sukarela melibatkan dirinya dalam kehidupan publik, termasuk siapa pun yang mencari kekuasaan, pengaruh, dan perhatian di publik.
4. Siapa pun yang terlibat dalam kontroversi publik, baik secara sukarela maupun tidak, misalnya terlibat dalam kasus-kasus yang mendapat perhatian media.

Secara sederhana, mereka biasanya adalah *public officials* (pejabat publik), *politicians* (politikus), *artists* (seniman, bintang film, musisi, selebgram, YouTuber), dan *sportsperson* (atlet). Tingkat privasi untuk empat profesi itu lebih rendah karena mereka dianggap secara sengaja mengekspos diri mereka ke depan publik dan mendapat keuntungan darinya. Warga biasa seperti seorang karyawan bank, misalnya, memiliki privasi yang sangat besar dan pengadilan cenderung memihak dirinya jika ada sengketa hukum terkait eksposur privasinya di media atau publik. Contoh sebaliknya, jika politikus menuntut pembuat konten dengan tuduhan mengganggu privasinya atau menghina, politikus itu memiliki beban lebih besar daripada karyawan itu untuk membuktikan adanya kerugian bagi dirinya atau niat jahat dari pembuat konten. Pajanan terhadap kehidupan pribadi si politikus dianggap sebagai konsekuensi dari profesi yang ia pilih.

Menurut penulis, cara pandang tersebut adil karena tentu saja privasi seorang politikus jauh lebih kecil daripada privasi seorang karyawan biasa.

Di Indonesia, akan lebih baik jika ada aturan atau kesepakatan terkait pembagian “kelas privasi” ini, sehingga tokoh publik menyadari bahwa profesinya memiliki konsekuensi berupa penurunan tingkat privasi yang bisa ia tuntut dalam kasus tuduhan pencemaran nama baik dan penghinaan.

D. Tata cara media sosial mengenali dan menentukan konten ilegal

Sejumlah riset (De Gregorio, 2020; Pirkova, 2022; dan Leong, 2022) menunjukkan bahwa algoritma dan *artificial intelligence* adalah tulang punggung pengaturan atau moderasi konten oleh media sosial, yang sangat mempengaruhi cara media sosial mengenali dan menentukan apakah suatu konten dianggap membahayakan pengguna.

Untuk Indonesia, selama ini penyelenggara media sosial belum cukup transparan dalam melakukan pengaturan konten bagi pengguna media sosial di Tanah Air. Misalnya, belum ada informasi terbuka mengenai kapan pengaturan dilakukan oleh *artificial intelligence* dan kapan oleh manusia, lalu apakah penyelenggara media sosial mempekerjakan warga Indonesia sebagai moderator konten dan berapa jumlahnya atau siapa mereka.

Sebagai gambaran global, Facebook menyatakan bahwa mereka menggunakan moderator manusia melalui penyediaan dari pihak ketiga yang berpusat di India, Irlandia, dan Filipina. Dalam praktiknya, moderator manusia yang dimiliki Facebook hanya mampu memahami kurang lebih 50 bahasa di dunia, padahal Facebook menyediakan layanan untuk lebih dari 100 bahasa. Hal ini memperjelas bahwa jumlah moderator manusia perlu ditingkatkan (Barrett, 2020). Bagi Indonesia, dengan ratusan bahasa daerah yang tersebar di lebih dari 13.000 pulau, tantangan itu jelas lebih besar dan perlu ada transparansi dari penyelenggara media sosial untuk memastikan pengaturan konten selaras dengan keberagaman Indonesia (Article 19, 2022).

Karena itu, PR2Media menilai perlu adanya regulasi yang mewajibkan penyelenggara media sosial menyampaikan mekanisme yang mereka gunakan untuk mengenali dan menentukan konten ilegal, baik melalui kecerdasan buatan, manusia, maupun kombinasi keduanya. Bagian D ini menyediakan pondasi bagi poin pengaturan di bagian-bagian di bawah ini. Usulan pasal untuk Bagian D ini bisa ditemukan dalam Lampiran 1 Kertas Kajian ini, yang disampaikan PR2Media kepada Panja Revisi UU ITE Komisi I DPR RI pada 23 Agustus 2023 (PR2Media, 2023).

E. Tata cara media sosial menanggapi aduan konten ilegal

Pemrosesan konten ilegal oleh penyelenggara media sosial pada dasarnya dilakukan secara proaktif oleh penyelenggara media sosial dan secara responsif berdasarkan aduan yang diterima dari pemerintah dan pengguna.

Secara *proaktif* berarti penyelenggara media sosial melalui moderator manusia dan/atau sarana otomatis melakukan fungsi moderasi konten. Fungsi ini seharusnya

merujuk pada aturan atau pedoman yang jelas dan tepat terkait dengan jenis konten apa yang dilarang, kapan, dan seperti apa tindakan akan diambil sehubungan dengan konten atau akun pengguna. Aturan dan kebijakan ini seharusnya mudah diakses oleh pengguna agar mereka memahami konten yang boleh dan tidak boleh hadir di dalam platform.

Secara *responsif* berarti penyelenggara media sosial mengambil tindakan terhadap konten atau akun tertentu atas dasar aduan yang disampaikan oleh pengguna dan pemerintah. Dalam hal ini, penting bagi penyelenggara media sosial melakukan evaluasi terhadap konten yang diadukan. Proses evaluasi seharusnya berlandaskan pada aturan atau pedoman yang jelas dan dilakukan secara transparan, tidak memihak, adil, dan menghargai hak asasi manusia.

Dengan mengacu pada sejumlah referensi, terutama Digital Services Act (European Commission, 2023) dan Santa Clara Principles² (Santa Clara Principles, 2021), PR2Media menyampaikan usulan regulasi sebagai berikut.

1. Mekanisme pengaduan bagi pengguna: Penyelenggara media sosial harus memiliki mekanisme bagi pengguna untuk mengadukan konten yang dianggap ilegal. Mekanisme ini harus mudah diakses dan mudah digunakan melalui sarana elektronik.
2. Mekanisme pemberitahuan bagi pengguna yang kontennya dianggap tidak mematuhi aturan: Mekanisme perlu dirancang sedemikian rupa untuk memfasilitasi penyampaian pemberitahuan yang jelas dan lengkap kepada pengguna yang kontennya dianggap tidak mematuhi regulasi maupun kebijakan *platform*. Untuk itu, penyelenggara media sosial harus mengambil langkah-langkah yang diperlukan untuk mengaktifkan dan memfasilitasi penyampaian pemberitahuan yang memuat elemen di bawah ini:
 - penjelasan tentang alasan mengapa informasi tersebut dipermasalahkan sebagai konten ilegal; termasuk aturan hukum yang dilanggar.
 - indikasi yang jelas tentang lokasi elektronik yang tepat dari informasi tersebut, seperti URL, dan, jika perlu, informasi tambahan yang memungkinkan identifikasi konten ilegal tersebut
 - pernyataan yang menegaskan keyakinan yang dapat dipercaya dari individu atau entitas yang mengirimkan pemberitahuan bahwa informasi dan dugaan yang terkandung di dalamnya adalah akurat dan lengkap.

² Tata cara yang disusun oleh organisasi hak asasi, pengacara, dan akademisi bagi platform digital yang melakukan pengaturan konten. Sejak kali pertama disuarakan dalam konferensi pengaturan konten di Santa Clara, California pada 2018, prinsip-prinsip ini telah didukung oleh 12 perusahaan digital besar, termasuk Meta, Google, Apple, dan Twitter.

3. Setelah ada aduan berisi informasi kontak elektronik dari individu atau entitas yang mengirimkannya, penyelenggara media sosial harus, tanpa penundaan yang tidak semestinya, mengirimkan konfirmasi penerimaan kepada individu atau entitas tersebut.
4. Penyelenggara media sosial harus memproses setiap aduan yang mereka terima berdasarkan mekanisme yang jelas dan mengambil keputusan sehubungan dengan pemberitahuan tersebut, secara tepat waktu, cermat, tidak sewenang-wenang dan objektif.
5. Apabila penyelenggara media sosial menggunakan sarana otomatis untuk pemrosesan atau pengambilan keputusan tersebut, mereka harus memasukkan informasi tentang penggunaan tersebut dalam pemberitahuan bagi pengguna yang kontennya dianggap tidak mematuhi aturan.
6. Sistem juga harus memfasilitasi pengadu untuk melacak proses aduan mereka.
7. Penyelenggara media sosial harus memberitahu pengadu tanpa penundaan yang tidak semestinya tentang keputusannya terkait dengan aduan tersebut, dan adanya mekanisme banding jika pengadu atau pengguna yang kontennya diadukan tidak puas dengan itu.

Mekanisme banding ini ada dua jenis, yaitu (1) untuk pengadu jika aduannya tidak dikabulkan, (2) untuk teradu yang kontennya dilaporkan dan diturunkan/ditandai.

8. Penyelenggara media sosial harus menjelaskan keputusannya dengan rinci, yang mencakup:
 - apakah akan menghapus atau menonaktifkan akses atau membatasi visibilitas informasi atau tidak;
 - apakah akan menanggukkan atau mengakhiri penyediaan layanan, seluruhnya atau sebagian, kepada pengguna;
 - apakah akan menanggukkan atau menghentikan akun pengguna atau tidak;
 - apakah akan menanggukkan, menghentikan, atau membatasi kemampuan untuk memonetisasi informasi yang diberikan oleh pengguna atau tidak.

Dengan penjelasan demikian, pengguna akan mengetahui dengan lebih pasti status konten atau akun mereka. Kejelasan ini juga membantu mereka dalam mengambil sikap sehubungan dengan keputusan penyelenggara platform.

Selain itu, sejumlah prinsip dalam Santa Clara Principles juga patut menjadi pertimbangan dalam penyusunan regulasi di Indonesia, yaitu:

Gambar 3. Prinsip Transparansi dan Akuntabilitas Penyelenggara Media Sosial

No.	Prinsip	Implementasi
<p>1. Hak Asasi Manusia dan Proses Hukum</p>	<p>Perusahaan media sosial harus memastikan bahwa hak asasi manusia dan pertimbangan proses hukum terintegrasi di semua tahapan proses moderasi konten, dan harus mempublikasikan informasi yang menguraikan bagaimana integrasi ini dibuat. Penyelenggara hanya boleh menggunakan proses otomatis untuk mengidentifikasi atau menghapus konten atau menanggihkan akun, baik dilengkapi dengan tinjauan manusia atau tidak, bila terdapat kepercayaan yang cukup tinggi terhadap kualitas dan akurasi proses tersebut. Penyelenggara juga harus memberi pengguna metode yang jelas dan dapat diakses untuk mendapatkan dukungan jika ada konten dan tindakan akun.</p>	<p>Pengguna harus diyakinkan bahwa pertimbangan hak asasi manusia telah diintegrasikan pada semua tahap proses moderasi konten, termasuk diberi tahu tentang:</p> <ul style="list-style-type: none"> • Bagaimana perusahaan mempertimbangkan hak asasi manusia—khususnya hak atas kebebasan berekspresi dan non-diskriminasi—dalam pengembangan aturan dan kebijakannya; • Bagaimana perusahaan mempertimbangkan pentingnya proses hukum saat menegakkan aturan dan kebijakannya, dan khususnya bagaimana proses tersebut memiliki integritas dan dikelola secara adil; • Sejauh mana perusahaan menggunakan proses otomatis dalam moderasi konten dan bagaimana perusahaan mempertimbangkan hak asasi manusia dalam penggunaan tersebut.
<p>2. Aturan dan Kebijakan yang Dapat Dimengerti</p>	<p>Perusahaan harus menerbitkan aturan dan kebijakan yang jelas dan tepat</p>	<p>Pengguna harus dapat dengan mudah memahami hal-hal berikut:</p> <ul style="list-style-type: none"> • Jenis konten apa yang dilarang oleh perusahaan dan akan dihapus,

No.	Prinsip	Implementasi
	<p>terkait dengan kapan tindakan akan diambil sehubungan dengan konten atau akun pengguna.</p>	<p>dengan panduan terperinci dan contoh konten yang diizinkan dan tidak diizinkan;</p> <ul style="list-style-type: none"> • Jenis konten apa yang akan diambil tindakan oleh perusahaan selain penghapusan, seperti penurunan peringkat algoritmik, dengan panduan terperinci dan contoh pada setiap jenis konten dan tindakan; • Keadaan ketika perusahaan akan menanggihkan akun pengguna, baik secara permanen maupun sementara.
<p>3. Kompetensi Budaya</p>	<p>Kompetensi budaya mensyaratkan, antara lain, bahwa mereka yang membuat keputusan moderasi memahami bahasa, budaya, dan konteks politik dan sosial dari konten yang mereka moderasi. Perusahaan harus memastikan bahwa aturan dan kebijakan mereka, dan penegakannya, mempertimbangkan keragaman budaya dan konteks di mana platform dan layanan mereka tersedia dan digunakan, dan harus mempublikasikan informasi tentang bagaimana pertimbangan ini telah terintegrasi dalam kaitannya dengan semua prinsip operasional. Perusahaan harus memastikan bahwa laporan,</p>	<ul style="list-style-type: none"> • Pengguna harus memiliki akses ke aturan dan kebijakan serta mekanisme pemberitahuan, banding, dan pelaporan dalam bahasa atau dialek yang mereka gunakan. Keputusan moderasi dibuat oleh mereka yang terbiasa dengan bahasa atau dialek yang relevan; • Keputusan moderasi dibuat dengan kesadaran yang memadai tentang konteks regional atau budaya yang relevan; • Perusahaan akan melaporkan data yang menunjukkan kompetensi bahasa, wilayah, dan budaya mereka untuk pengguna yang mereka layani, seperti angka yang menunjukkan distribusi bahasa dan geografis dari moderator konten mereka.

No.	Prinsip	Implementasi
	<p>pemberitahuan, dan proses banding tersedia dalam bahasa yang digunakan pengguna untuk berinteraksi dengan layanan, dan bahwa pengguna tidak dirugikan selama proses moderasi konten berdasarkan bahasa, negara, atau wilayah.</p>	
<p>4. Keterlibatan Negara/ Pemerintah dalam Moderasi Konten</p>	<p>Perusahaan harus mengenali risiko tertentu terhadap hak pengguna yang dihasilkan dari keterlibatan negara dalam proses moderasi konten. Ini termasuk keterlibatan negara dalam pengembangan dan penegakan peraturan dan kebijakan perusahaan, baik untuk mematuhi undang-undang setempat atau melayani kepentingan negara lainnya. Perhatian khusus diajukan oleh tuntutan dan permintaan dari aktor negara (termasuk badan pemerintah, otoritas pengatur, lembaga penegak hukum, dan pengadilan) untuk penghapusan konten atau penangguhan akun.</p>	<ul style="list-style-type: none"> • Pengguna harus mengetahui kapan aktor negara telah meminta atau berpartisipasi dalam tindakan apa pun pada konten atau akun mereka. Pengguna juga harus mengetahui apakah perusahaan yakin bahwa tindakan tersebut diwajibkan oleh undang-undang yang relevan. Sementara beberapa perusahaan sekarang melaporkan tuntutan negara untuk pembatasan konten berdasarkan undang-undang sebagai bagian dari pelaporan transparansi mereka, keterlibatan negara lain tidak dilaporkan secara publik maupun kepada pengguna yang terkena tindakan. Tetapi perusahaan harus secara jelas melaporkan kepada pengguna ketika ada keterlibatan negara dalam penegakan peraturan dan kebijakan perusahaan. • Secara khusus, pengguna harus dapat mengakses: Detail peraturan atau kebijakan apa pun, baik yang berlaku secara global atau di yurisdiksi tertentu, yang berupaya mencerminkan persyaratan hukum setempat.

No.	Prinsip	Implementasi
		<ul style="list-style-type: none"> • Perincian hubungan kerja formal atau informal dan/atau perjanjian yang dimiliki perusahaan dengan aktor negara dalam hal menandai konten atau akun atau tindakan lain yang diambil oleh perusahaan. • Detail proses penilaian konten atau akun yang ditandai oleh aktor negara, baik berdasarkan aturan atau kebijakan perusahaan atau undang-undang setempat. • Detail permintaan status ke pos tindakan dan akun.
5.	<p>Integritas dan Kejelasan</p> <p>Perusahaan harus memastikan bahwa sistem moderasi kontennya, termasuk komponen otomatis dan non-otomatis, bekerja dengan andal dan efektif. Ini termasuk mengupayakan akurasi dan non diskriminasi dalam metode deteksi, mengirimkan ke penilaian reguler, dan menyediakan mekanisme pemberitahuan dan banding secara adil. Perusahaan harus secara aktif memantau kualitas pengambilan keputusan mereka untuk memastikan tingkat kepercayaan yang tinggi, dan didorong untuk berbagi data secara publik tentang keakuratan sistem mereka dan untuk</p>	<p>Pengguna harus yakin bahwa keputusan tentang konten mereka dibuat dengan sangat hati-hati dan menghormati hak asasi manusia. Pengguna harus mengetahui kapan keputusan moderasi konten dibuat atau dibantu oleh alat otomatis, dan memiliki pemahaman tingkat tinggi tentang logika pengambilan keputusan yang digunakan dalam proses otomatis terkait konten. Perusahaan juga harus secara jelas menguraikan kontrol apa yang dapat diakses pengguna yang memungkinkan mereka mengelola cara konten mereka dikuratori dengan sistem algoritma, dan apa dampak kontrol ini terhadap pengalaman daring pengguna.</p>

No.	Prinsip	Implementasi
	<p>membuka proses dan sistem algoritmik mereka untuk audit eksternal berkala. Perusahaan harus bekerja untuk memastikan bahwa permintaan tindakan adalah asli dan bukan hasil dari bot atau serangan terkoordinasi.</p>	

Sumber: Santa Clara Principles (2021).

F. Media sosial menyediakan mekanisme banding

Pemrosesan konten ilegal secara proaktif maupun responsif tidak selalu memuaskan semua pihak. Dalam kondisi tertentu mungkin saja pengguna dan pemerintah mengajukan protes atas tindakan platform karena merasa tidak puas. Untuk itu, penyelenggara media sosial berkewajiban menyediakan sistem yang memungkinkan pengguna atau pemerintah mengajukan aduan/banding.

Namun, proses banding kepada penyelenggara media sosial sulit dilakukan akibat minimnya informasi terkait tata cara yang harus dilalui oleh pengguna. Berdasarkan wawancara PR2Media dengan pengguna yang akunnya pernah ditangguhkan, pengguna mengakui bahwa akun miliknya dapat kembali diakses akibat bantuan dari orang yang ia kenal di platform tersebut, tanpa ia menerima respons atau notifikasi langsung dari pihak platform (Wendratama et al., 2023). Sebagian besar pengguna media sosial di Indonesia tidak mengetahui cara pasti untuk mengajukan banding dengan pihak platform. Sementara itu, kelompok masyarakat sipil terkemuka dalam kapasitasnya sebagai mitra resmi platform, sering kali merasa tidak berdaya dalam proses negosiasi dengan platform (Article 19, 2022). Oleh karenanya, permasalahan ini cenderung dikaitkan dengan pembatasan kebebasan berekspresi dan keselamatan individu dan publik.

Adanya banding dari pengguna perlu ditanggapi serius oleh platform. Tanggapan ini merupakan bentuk tanggung jawab platform kepada warga masyarakat atas pelayanan yang mereka ciptakan. Tanggapan ini juga memiliki arti penting untuk mencegah tindakan kesewenang-wenangan, yang dapat mengancam kebebasan berekspresi dan

berpendapat warga. Berdasarkan rekomendasi penelitian CfDS UGM (Rahman et al., 2022), mekanisme banding untuk platform media sosial menjadi salah satu hal yang harus diatur oleh negara, melalui adanya reformasi terlebih dahulu terhadap mekanisme penanganan di UU ITE. Oleh sebab itu, perlu ada regulasi atau kebijakan yang mengatur suatu tata cara bagaimana platform seharusnya merespons banding.

Digital Services Act (DSA) mewajibkan VLOPs (*very large online platform*), yang memiliki 10% pengguna aktif bulanan di Uni Eropa, untuk menyediakan tiga pilihan mekanisme banding bagi pengguna. Tiga pilihan ini berlaku untuk mekanisme banding bagi pengadu yang aduannya tidak dikabulkan maupun bagi teradu yang kontennya diturunkan/ditandai.

1. Mekanisme banding internal

Ketersediaan akses atau medium untuk menyampaikan banding atas keputusan platform merupakan hal penting bagi pengguna yang merasa kebebasan berekspresinya dilanggar oleh platform atau merasa platform membuat keputusan yang tidak tepat.

Penyelenggara platform, dalam hal ini penyelenggara media sosial, harus memastikan bahwa sistem penanganan pengaduan internal mereka mudah diakses, mudah digunakan, dan memfasilitasi pengajuan pengaduan yang cukup tepat dan didukung bukti yang memadai.

Sistem juga selayaknya memberikan kesempatan bagi pengguna (pengadu maupun teradu) untuk menyajikan informasi tambahan untuk mendukung banding mereka yang akan dipertimbangkan dalam peninjauan.

Hal penting yang perlu dipatuhi oleh penyelenggara platform dalam menangani banding ini adalah ketepatan waktu memberikan respons. Hasil peninjauan dan keputusan dari mekanisme banding harus dilakukan secara cepat jika konten yang diadukan bersifat sensitif terhadap waktu, seperti konten politik selama periode pemilihan.

Penanganan juga seharusnya tidak diskriminatif, cermat dan tidak sewenang-wenang. Ketentuan ini senada dengan prinsip-prinsip yang terangkum dalam Santa Clara Principles, bahwa dalam moderasi konten haruslah transparan dan akuntabel dengan memperhatikan keadilan, tidak memihak, proporsional, dan menghormati hak pengguna. Proporsional berarti penyelenggara media sosial harus memprioritaskan pemberian banding untuk pembatasan yang paling berat, seperti penghapusan konten dan penangguhan akun.

Penyelenggara platform harus memastikan bahwa keputusan diambil di bawah pengawasan staf yang memenuhi syarat, dan tidak semata-mata atas dasar sarana

otomatis. Staf di sini seharusnya memiliki kompetensi kultural. Sebagaimana disebutkan dalam Santa Clara Principles, mereka yang membuat keputusan moderasi haruslah memahami bahasa, budaya, dan konteks politik dan sosial dari postingan yang mereka moderasi.

Dengan mengacu pada Santa Clara Principles, PR2Media mengusulkan mekanisme banding yang disediakan penyelenggara media sosial perlu mencakup hal-hal sebagai berikut:

- Proses yang jelas dan mudah diakses oleh pengguna, yang mencakup detail waktu yang harus tertulis secara jelas. Hal ini ditujukan untuk melacak proses kemajuan aduan pengguna.
- Tinjauan atau penilaian dari orang yang tidak terlibat dalam keputusan awal, sehingga bisa memberikan *second opinion*.
- Pemahaman bahasa dan konteks budaya yang dimiliki oleh orang yang terlibat dalam proses peninjauan karena pihak platform media sosial cenderung belum menunjukkan pemahaman yang memadai tentang kompleksitas konteks lokal (Article 19, 2022).
- Kesempatan bagi pengguna untuk menyertakan bukti tambahan untuk mendukung proses banding
- Pemberitahuan hasil peninjauan disertai dengan alasan dan pertimbangan yang cukup jelas, agar pengguna memahami keputusan tersebut.

2. Mekanisme eksternal di luar pengadilan

Dalam kajian ini, PR2Media mengusulkan adanya Dewan Media Sosial, yang akan melayani mekanisme eksternal tersebut. Keberadaan dan fungsi Dewan Media Sosial akan diuraikan dalam bagian berikutnya.

3. Proses peradilan

DSA menyatakan bahwa pengguna juga bisa menempuh proses peradilan terkait keputusan moderasi konten dari penyelenggara media sosial. Jika peradilan memutuskan bahwa keputusan penyelenggara membuat tidak bisa diaksesnya sebuah konten tidak sesuai dengan aturan (*terms and conditions*) yang dibuat penyelenggara media sosial, maka penyelenggara media sosial wajib memulihkan konten tersebut.

G. Dewan Media Sosial

Tim perumus kajian ini dan masyarakat sipil yang terlibat dalam proses riset untuk advokasi kebijakan mendorong agar negara Indonesia mulai mengadopsi model lembaga independen dalam mengawasi dan memberikan rekomendasi terkait cara penyelenggara media sosial melakukan moderasi konten. Independen artinya terpisah dari unsur kepentingan langsung/tidak langsung atau keterwakilan dari pemerintah dan perusahaan media sosial.

Kami menilai, gagasan soal Social Media Council atau Dewan Media Sosial yang bergema di lembaga-lembaga internasional dapat menjadi rujukan model kelembagaan. Pada 2019, Inkubator Kebijakan Digital Global (GDPI) Stanford University, ARTICLE 19, dan Pelapor Khusus PBB untuk Hak atas Kebebasan Berpendapat dan Bereksprei, David Kaye, (Article 19, 2021) mengadakan pertemuan untuk membahas solusi untuk mengatasi tantangan konten ilegal dan mengusulkan pembentukan *multi stakeholder agency*: Dewan Media Sosial (DMS).

Tantangan utama moderasi konten adalah: Bagaimana kita bisa menyeimbangkan tanggung jawab untuk melindungi kebebasan berbicara dengan kebutuhan untuk mencegah efek berbahaya, sambil mempertimbangkan tantangan yang ditimbulkan oleh privatisasi ruang digital di mana bentuk-bentuk ucapan ini muncul?

Dewan Media Sosial adalah sebuah mekanisme akuntabilitas multi-pihak yang akan menyediakan forum yang terbuka, transparan, independen, dan akuntabel untuk mengatasi masalah moderasi konten di platform media sosial berdasarkan standar internasional tentang hak asasi manusia. Model DMS mengedepankan pendekatan sukarela untuk pengawasan moderasi konten: peserta yang berasal dari penyelenggara media sosial, pemerintah, pakar, dan kelompok masyarakat sipil) mendaftar untuk menjadi bagian dari lembaga ini. Kekuatan dan efisiensinya bergantung pada kepatuhan sukarela oleh platform, yang saat melakukan registrasi, berkomitmen menghormati dan melaksanakan keputusan (atau rekomendasi) DMS dengan itikad baik. Insentif bagi penyelenggara media sosial untuk terlibat di dalam DMS muncul dari kebutuhan mereka untuk memulihkan kepercayaan pengguna terhadap transparansi dan akuntabilitas tata kelola mereka serta partisipasi perwakilan pemerintah di dalam DMS yang memberikan legitimasi terhadap lembaga tersebut.

Untuk konteks Indonesia, PR2Media mengusulkan DMS memiliki tiga peran utama yaitu:

- Memberikan panduan umum tentang praktik moderasi konten untuk memastikan praktik ini mengikuti standar internasional tentang kebebasan bereksprei dan

nilai-nilai yang berkembang di Indonesia (konteks lokal, mengingat keberagaman Indonesia).

- Menjadi sebuah forum bagi berbagai pemangku kepentingan mendiskusikan rekomendasi terkait moderasi konten.
- Menjadi lembaga banding administratif bagi pengguna yang tidak puas terhadap keputusan moderasi konten oleh penyelenggara media sosial. Selain menerima banding, DMS juga mengulas keputusan-keputusan moderasi konten yang dibuat penyelenggara media sosial berdasarkan laporan yang mereka buat.

Perumus melihat model DMS ini tidak hanya perlu diterapkan di Jakarta, tetapi juga memiliki jaringan DMS di setiap provinsi (karena keberagaman Indonesia) bahkan internasional untuk bisa memenuhi tiga peran itu dengan baik. Dengan demikian, forum multi-pihak, transparan, akuntabel, dan independen ini bisa menenun kebebasan berekspresi dalam semua aspek moderasi dan distribusi konten platform media sosial, mulai dari mengintegrasikan standar internasional dalam keputusan untuk menghapus atau menurunkan konten, hingga memastikan paparan terhadap keragaman informasi dan ide melalui distribusi algoritma.

Lembaga multipihak ini (terdiri dari perwakilan platform media sosial, pemerintah, peneliti/akademisi, organisasi masyarakat sipil, tokoh agama, ahli hukum, dan lain-lain) menjadi pilihan yang tepat untuk mengawasi dan menilai kepatuhan penyelenggara media sosial terhadap regulasi.

Sebagai lembaga independen, DMS berbentuk lembaga non-struktural dan bertanggung jawab kepada presiden.

DMS akan mengambil pendekatan yang berbeda dari dewan pengawas internal yang ada di platform, karena DMS merupakan forum multi-pihak. Dewan media sosial akan menyertakan perwakilan dari kelompok terpinggirkan dan organisasi masyarakat sipil di seluruh spektrum politik lokal dan nasional untuk memastikan bahwa rekomendasinya menggabungkan pengalaman dan saran dari orang-orang yang paling sering terpengaruh oleh tindakan moderasi konten. Dengan mengamatkan pertemuan rutin dan berbagi informasi secara terbuka dan partisipatif, DMS dapat menjadi institusi yang kuat untuk mengatasi masalah ketidakpercayaan publik kepada platform media sosial.

DMS juga perlu dilihat sebagai alternatif solusi terhadap dua isu, yaitu 1) *self-regulatory* platform yang selama ini kurang efektif mengatasi konten ilegal, 2) regulasi dan implementasi regulasi di Indonesia yang belum efektif mengatasi konten ilegal dan bahkan cenderung represif.

Poin nomor dua di atas terkait dengan *due process of law* yang perlu ditempuh pemerintah dalam menjalankan wewenangnya memutus akses ke penyelenggara sistem elektronik, sebagaimana diatur dalam Pasal 40 ayat 2b Undang-Undang Informasi dan Transaksi Elektronik.

Pasal 40 ayat 2b berbunyi, “Dalam melakukan pencegahan sebagaimana dimaksud pada ayat (2a), Pemerintah berwenang melakukan pemutusan akses dan/atau memerintahkan kepada Penyelenggara Sistem Elektronik untuk melakukan pemutusan akses terhadap Informasi Elektronik dan/atau Dokumen Elektronik yang memiliki muatan yang melanggar hukum.”

Ayat tersebut membuat mekanisme pemutusan akses sepenuhnya berada di tangan pemerintah, alias tidak ada mekanisme pengawasan dari lembaga lain terkait tindakan pemutusan akses. Dalam kondisi ini, keberadaan DMS bisa menjadi sebuah solusi untuk mengurangi risiko tindakan yang represif dari pemerintah.

Sumber Pembiayaan Dewan Media Sosial

Pengaturan DMS sebagai lembaga independen ditetapkan oleh undang-undang yang mengatur penyebaran konten ilegal di platform digital. Karena amanat undang-undang, pendanaan DMS juga berasal dari negara, di samping iuran dari konstituen (yaitu penyelenggara media sosial). Model pendanaan seperti ini juga telah dilakukan oleh Dewan Pers, yang pembiayaannya berasal dari perusahaan pers, organisasi pers, bantuan dari negara, dan bantuan lain yang tidak mengikat (Undang-Undang No. 40 Tahun 1999 tentang Pers).

DMS tidak perlu mengurus kebijakan pajak atau persaingan usaha, karena ini sudah masuk ranah bisnis. Dewan menetapkan standar moderasi konten yang adil, andal, transparan, dan tidak sewenang-wenang. Pada saat keputusan moderasi konten oleh penyelenggara media sosial semakin ditentukan oleh kepentingan bisnis mereka, DMS menawarkan metode yang relatif cepat untuk mengoordinasikan dan mengatasi masalah akuntabilitas tata kelola konten yang mendesak.

Akhirnya, keberadaan DMS dan usulan regulasi tata kelola konten/moderasinya di media sosial dalam dokumen ini secara teoretis sudah selaras dengan dua pendekatan

kebijakan berikut. Pertama, quasi-regulasi: Pemerintah mendorong bisnis digital bersidang dan menetapkan peraturan dan mekanisme kepatuhan manajemen konten. Namun, pemerintah tidak menentukan sifat aturan tersebut atau penegakannya. Kedua, co-regulation: industri menciptakan dan mengelola peraturan dan standarnya sendiri. Namun, pemerintah menetapkan undang-undang untuk menegakkan pengaturan tersebut. Media penyiaran adalah contoh klasik pendekatan ini. Industri media sosial ingin mengatur sendiri, publik ingin ada pengaturan bersama.

H. Laporan tahunan dari media sosial

Indonesia memiliki tiga regulasi mengenai aturan konten yang juga mengatur platform media sosial, yaitu Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggara Sistem dan Transaksi Elektronik (PP 71/2019), dan Peraturan Menteri Komunikasi dan Informatika Nomor 5 Tahun 2020 (Permenkominfo 5/2020).

Dalam menangani konten ilegal dan berbahaya, regulasi tersebut masih belum mengatur tata cara pelaporan hasil praktik moderasi konten, yang berhubungan dengan transparansi dan akuntabilitas dari platform media sosial terhadap pengguna.

Informasi tentang praktik moderasi konten yang dilakukan oleh platform media sosial cenderung sulit ditemukan. Maka dari itu, penjelasan mengenai langkah dan cara yang dilakukan platform media sosial ketika melakukan moderasi, minim diketahui oleh masyarakat umum (Rahman dkk., 2022).

Berdasarkan wawancara PR2Media terhadap Yendra Budiana dan Ajiwan Arief dari kelompok minoritas, yaitu Ahmadiyah dan SIGAB Indonesia (Wendratama et al., 2023), beberapa konten yang pernah mereka coba laporkan, berakhir lamban dan tidak direspons oleh platform media sosial. Padahal, konten ilegal yang dilaporkan memiliki dampak buruk bagi kelompok mereka karena tersebar secara masif, sehingga butuh respons dan proses penghapusan konten secara cepat dari pihak platform. Damar Juniarto dari SAFEnet, mengatakan bahwa penanganan konten ilegal dan berbahaya oleh platform masih kurang memuaskan, akibat platform yang kurang responsif dan transparansi yang disediakan hanya berbasis angka tanpa adanya pola penyelesaian yang tidak dijelaskan (Wendratama et al., 2023).

Maka dari itu, platform media sosial diharapkan menyediakan laporan untuk mendukung transparansi dan akuntabilitas mereka terhadap pengguna. Merujuk pada Digital Services Act (DSA), pemerintah Uni Eropa mewajibkan VLOPs (*very large online platforms*), seperti Facebook, Instagram, YouTube, dan TikTok, untuk menerbitkan laporan setiap enam bulan sekali (Nosák, 2021). Sedangkan, pemerintah Indonesia

belum mewajibkan platform media sosial untuk menyerahkan laporan terkait praktik moderasi yang sudah dilakukan per tahunnya. Untuk mengawasi kinerja platform *online* atau Penyelenggara Sistem Elektronik (PSE) lingkup privat, Kemenkominfo masih dalam tahap mewajibkan mekanisme pendaftaran bagi PSE dan menjalankan mekanisme pengendalian terhadap PSE sesuai dengan Permenkominfo 5/2020 (Rizkinaswara, 2021)

Oleh karena itu, PR2Media mengusulkan beberapa poin yang perlu dimuat dalam laporan tahunan platform media sosial, yaitu:

Pertama, penyelenggara media sosial harus menyampaikan secara terbuka tentang tata cara dalam mengenali dan menentukan informasi elektronik dan/atau dokumen elektronik yang melanggar hukum, baik yang dilakukan oleh sistem otomatis (*artificial intelligence*) maupun oleh orang yang dipekerjakan (moderator manusia). Walaupun *artificial intelligence* (AI) mampu bekerja lebih cepat dibandingkan moderator manusia, tetapi AI masih mengalami kesulitan untuk memperhitungkan konteks, sarkasme, dan makna kultural (Duerte dkk., 2017).

Dalam poin ini diharapkan penyampaian pelaporan dapat dikemas secara sederhana dengan detail menyeluruh mengenai proses moderasi dan kebijakan yang dilanggar (West, 2018). Informasi mengenai jumlah moderator manusia yang dipekerjakan pun hanya disampaikan oleh beberapa platform secara global, seperti Facebook yang mempekerjakan 15.000 moderator manusia (Koetsier, 2020), serta Youtube dan Google sebanyak 10.000 orang moderator (Newton, 2019). Sayangnya, informasi spesifik mengenai jumlah moderator manusia yang dipekerjakan di Indonesia belum pernah disediakan oleh platform.

Format pelaporan diharapkan menyediakan informasi kuantitatif dan kualitatif tentang hasil aktual dan dampak yang dihasilkan dari moderasi konten berdasarkan ketentuan layanan atau sesuai permintaan pelapor, serta berdasarkan hasil moderasi dari sarana otomatis/AI atau moderator manusia (UNESCO, 2023).

Kedua, laporan tahunan ini menginformasikan jumlah aduan yang diterima dan tindakan yang dilakukan untuk menangani aduan tersebut. Merujuk pada Santa Clara Principles (2021) yang sudah menjalani proses konsultasi dengan lebih dari 50 organisasi dan individu, platform media sosial diharapkan menginformasikan tentang konten dan akun yang dikenai sanksi oleh platform terkait hal-hal berikut:

- Jumlah total konten yang dikenai sanksi dan jumlah total akun ditangguhkan.
- Jumlah banding terhadap keputusan sanksi terhadap konten atau akun.
- Jumlah (atau persentase) banding yang berhasil sehingga konten atau akun dipulihkan, dan jumlah (atau persentase) banding yang gagal;
- Jumlah konten atau akun yang dipulihkan oleh perusahaan secara proaktif, tanpa banding apa pun, setelah mengetahui bahwa sanksi terhadap konten atau akun tersebut tidak tepat.

- Angka terkait penghapusan dan pembatasan konten yang dilakukan selama periode krisis, seperti selama pandemi COVID-19 dan periode konflik kekerasan.

Sedangkan, untuk tindakan penanganan yang dilakukan, Pirkova (2022) menyampaikan bahwa DSA mewajibkan semua platform untuk melaporkan secara publik tentang cara penggunaan alat moderasi konten otomatis, tingkat kesalahan alat, dan informasi tentang pelatihan dan bantuan yang diterapkan oleh platform kepada moderator konten yang dimiliki. Untuk VLOPs (*very large online platforms*), DSA menetapkan adanya laporan mengenai penilaian risiko yang dibawa platform kepada pengguna, tindakan mitigasi risiko, serta laporan audit dan pelaksanaan audit (Nosák, 2021). Masyarakat berhak mendapatkan transparansi terkait penanganan konten yang mereka adukan, agar platform dapat memberikan alasan yang jelas terkait penilaian mereka terhadap konten yang diadukan, serta secara tidak langsung menjaga hubungan afektif antara pengguna dan platform media sosial. Tidak hanya bagi masyarakat, menurut UNESCO (2023), transparansi yang bermakna akan berguna untuk informasi bagi pemangku kepentingan dalam membuat keputusan yang tepat. Oleh karenanya, laporan tahunan diharapkan bisa disampaikan dengan efektif dan mudah dipahami.

Untuk memudahkan, berikut gambaran rekomendasi dari PR2Media terkait hal yang perlu tersedia di dalam laporan tahunan, yaitu:

Gambar 4. Isi Laporan Tahunan

No.	Bagian	Penjelasan
1.	Jumlah dan jenis konten yang diadukan	<ul style="list-style-type: none"> • Jumlah total konten yang dikenai sanksi dan jumlah total akun ditangguhkan. • Jumlah banding terhadap keputusan sanksi terhadap konten atau akun. • Jumlah (atau persentase) banding yang berhasil sehingga konten atau akun dipulihkan, dan jumlah (atau persentase) banding yang gagal. • Jumlah konten atau akun yang dipulihkan oleh perusahaan secara proaktif, tanpa banding apa pun, setelah mengetahui bahwa sanksi terhadap konten atau akun tersebut tidak tepat. • Angka terkait penghapusan dan pembatasan konten yang dilakukan selama periode krisis, seperti selama pandemi COVID-19 dan periode konflik kekerasan. <p>Acuan: Santa Clara Principles (2021)</p>

No.	Bagian	Penjelasan
2.	Pihak pelapor	<ul style="list-style-type: none"> • Pemerintah • <i>Trusted flagger</i> – organisasi yang memiliki kompetensi dalam mengenali dan menandai konten ilegal dan berbahaya (Wendratama et al., 2023) • Pengguna
3.	Cara platform media sosial menindaklanjuti aduan	<ul style="list-style-type: none"> • Sistem otomatis (<i>artificial intelligence/AI</i>) <ul style="list-style-type: none"> - Tata cara penggunaan alat, yang dapat dilengkapi dengan informasi alat yang digunakan dalam tindakan tertentu - Tingkat kesalahan alat, termasuk tingkat kepercayaan platform media sosial terhadap alat yang digunakan • Moderator manusia <ul style="list-style-type: none"> Pelatihan dan bantuan yang diterapkan oleh platform <p>Acuan: Digital Services Act (dalam Pirkova, 2022) dan Santa Clara Principles (2021)</p> <hr/> <ul style="list-style-type: none"> • Penilaian risiko platform • Tindakan mitigasi risiko • Laporan audit dan pelaksanaan audit <p>Acuan: Digital Services Act, khusus untuk VLOPs (dalam Nosák, 2021)</p>

I. Audit independen terhadap media sosial

Untuk menilai kepatuhan penyelenggara media sosial terhadap regulasi ini, perlu ada audit independen terhadap tata kelola penyelenggara media sosial. Audit ini berfokus pada kepatuhan terhadap regulasi, tidak mencakup audit keuangan.

Aspek yang perlu diaudit antara lain mekanisme kurasi konten (misalnya bagaimana konten tertentu direkomendasikan kepada pengguna tertentu oleh sistem otomatis), mekanisme moderasi konten (misalnya bagaimana konten yang melanggar ketentuan platform dikenali dan ditandai atau diturunkan oleh sistem otomatis dan keterlibatan moderator manusia), efektivitas mekanisme banding, dan mekanisme manajemen risiko.

Audit ini perlu dilakukan oleh organisasi (auditor) eksternal yang terbukti memiliki keahlian profesional dalam manajemen risiko dan kompetensi teknis, serta memiliki

rekam jejak melakukan audit independen. Audit ini dilakukan minimal satu sekali setahun dan dibiayai oleh penyelenggara media sosial.

Di bawah ini adalah sejumlah poin pelaksanaan audit menurut DSA (European Commission, 2023), yang juga bisa diadaptasi untuk konteks Indonesia.

1. Penyelenggara media sosial memberikan kerjasama dan bantuan yang diperlukan kepada organisasi yang melaksanakan audit, termasuk dengan memberikan akses ke auditor terhadap semua data yang diperlukan untuk melaksanakan audit dengan benar, termasuk, jika relevan, data yang terkait dengan sistem algoritma.
2. Audit harus dilakukan sesuai dengan praktik industri terbaik dan etika profesional yang tinggi serta objektivitas, dengan memperhatikan, standar dan kode praktik audit.
3. Auditor harus menjamin kerahasiaan, keamanan dan keutuhan informasi, seperti rahasia perusahaan, yang diperolehnya dalam melaksanakan tugasnya. Jaminan ini tidak boleh menjadi sarana untuk mengelak dari penerapan kewajiban audit.
4. Laporan audit harus bisa dibuktikan supaya bisa memberikan penjelasan yang berarti tentang kegiatan yang dilakukan dan kesimpulan yang dicapai. Laporan audit harus membantu menginformasikan, dan jika perlu, memberikan rekomendasi tindakan yang perlu diambil oleh penyelenggara media sosial untuk mematuhi regulasi ini.
5. Laporan audit harus dikirimkan kepada lembaga yang berwenang (pemerintah dan Dewan Media Sosial).
6. Penyelenggara media sosial harus mengirimkan laporan kepada lembaga yang berwenang (pemerintah dan Dewan Media Sosial) tentang bagaimana mereka menanggapi laporan audit tersebut.

Di Uni Eropa, pelaksanaan audit independen terhadap platform besar (dengan jumlah pengguna di Uni Eropa lebih dari 45 juta orang) baru dimulai pada akhir Agustus 2023 (Deloitte Legal, 2023), sehingga belum pernah ada laporan audit independen yang diterbitkan. Tenggat untuk laporan audit pertama adalah Agustus 2024 dan dipublikasikan oleh platform besar paling lambat November 2024 (Tremau, 2023). Saat ini belum ada auditor independen yang dinyatakan melakukan audit tersebut. Meski demikian, sejumlah perusahaan audit sudah menyatakan kesiapannya melakukan audit ketaatan platform besar terhadap DSA, antara lain Deloitte (Cankett & Fackovcova, 2022) dan Ernst & Young (Legat & Guzy, 2023).

Penyampaian kertas kerja ini kepada pengambil kebijakan di Indonesia (Pemerintah dan DPR RI) dan penyebarluasannya kepada publik merupakan bentuk partisipasi PR2Media dalam proses reformasi regulasi terkait tata kelola media digital di Indonesia. Dengan publikasi ini, diharapkan terjadi dialog terbuka dan dokumen ini akan menjadi catatan sejarah partisipasi publik dalam pengaturan media sosial yang selaras dengan kepentingan umum. Publikasi ini juga disemangati oleh kerja-kerja advokasi PR2Media yang menempatkan ruang digital (internet dan media sosial) sebagai hak konstitusi warga (*Internet Constitutionalism*).

Kami meyakini, media sosial merupakan institusi di ruang publik yang menyediakan informasi dan pengetahuan secara terbuka sebagai kebutuhan dasar (*public domain*), sehingga ia patut diatur. Dalam konteks ini, semua pihak (terutama pemerintah dan platform media sosial) mutlak memiliki regulasi yang menghormati norma-norma HAM, seperti konvensi internasional terkait kebebasan berekspresi yang sudah diratifikasi oleh Indonesia, yaitu International Covenant on Civil and Political Rights (ICCPR, 1966) dan International Convention on the Elimination of All Forms of Racial Discrimination (ICERD, 1965).

Kovenan Internasional tentang Hak-Hak Sipil dan Politik (ICCPR) adalah perjanjian multilateral yang mengikat negara-negara untuk menghormati hak-hak sipil dan politik individu, termasuk hak untuk hidup, kebebasan beragama, kebebasan berbicara, kebebasan berkumpul, hak pemilihan, dan hak atas proses hukum dan pengadilan yang adil. Perjanjian ini sudah diratifikasi oleh Indonesia pada 2005 (antara lain melalui Undang-Undang No. 12 Tahun 2005 tentang Pengesahan *International Covenant on Civil and Political Rights*).

Sementara itu, Konvensi Internasional tentang Penghapusan Segala Bentuk Diskriminasi Rasial (ICERD) adalah konvensi Perserikatan Bangsa-Bangsa yang mengikat negara anggotanya untuk menghapus diskriminasi rasial dan memajukan pemahaman di antara semua ras. Konvensi ini juga mewajibkan para pihak untuk mengkriminalisasi ujaran kebencian dan mengkriminalkan keanggotaan dalam organisasi rasis. Konvensi ini sudah diratifikasi oleh Indonesia pada 1999 (Undang-Undang No. 29 Tahun 1999

tentang Pengesahan International Convention on The Elimination of All Forms of Racial Discrimination 1965 (Konvensi Internasional tentang Penghapusan Segala Bentuk Diskriminasi Rasial 1965, 1999).

Terakhir, kerangka kerja dan undang-undang baru yang diusulkan (sebagai hasil dari revisi total UU ITE) harus memastikan kerja peninjauan regulasi secara berkala untuk memastikan bahwa peraturan dan pedoman tersebut mutakhir, efektif, dan tidak terlalu memberatkan. Peninjauan berkala ini juga perlu memuat tata cara pengumpulan bukti penerapan regulasi dan dampaknya atas mitigasi penyebaran konten ilegal di media sosial. Dengan demikian, regulasi selalu bisa menjawab tantangan digital terkini yang sangat dinamis.

Kami mengucapkan terima kasih kepada semua pihak yang terlibat membantu terwujudnya kertas kerja ini. Mereka antara lain: tim perumus yang merupakan perwakilan dari masyarakat sipil pegiat hak asasi di bidang media seperti SAFEnet, LBH Pers, Mafindo, ICJR, dan AJI Indonesia. Terima kasih secara khusus kami sampaikan kepada Yayasan Tifa, yang memberi dukungan pendanaan. Selanjutnya, kertas kebijakan ini menjadi milik publik dan kami membuka diri untuk masukan dan koreksi untuk perbaikannya.

Buku, Jurnal, Artikel, dan Video

- Aichner, T., Grünfelder, M., Maurer, O., & Jegeni, D. (2021). Twenty-five years of social media: A review of social media applications and definitions from 1994 to 2019. *Cyberpsychology, Behavior, and Social Networking*, 24(4), 215–222. <https://doi.org/10.1089/cyber.2020.0134>
- Article 19. (2022). *Content moderation and local stakeholders in Indonesia*. Article 19. <https://www.article19.org/wp-content/uploads/2022/06/Indonesia-country-report.pdf>
- Article 19. (2021, October 12). Social Media Councils: One piece in the puzzle of content moderation. *Article 19*. <https://www.article19.org/resources/social-media-councils-moderation/>
- Article 20 Digital Services Act. (n.d.). Article 20, internal complaint-handling system - the Digital Services Act (DSA). *Cyber Risk GmbH*. https://www.eu-digital-services-act.com/Digital_Services_Act_Article_20.html
- Barrett, P. M. (2020). *Who moderates the social media giants? A call to end outsourcing*. New York University Stern Center for Business and Human Rights. https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/5ed9854bf618c710cb55be98/1591313740497/NYU+Content+Moderation+Report_June+8+2020.pdf
- Cankett, M., & Fackovcova, L. (2022, May 18). EU Digital Services Act: Are you ready for audit?. *Deloitte*. <https://www2.deloitte.com/uk/en/blog/auditandassurance/2022/eu-digital-services-act-are-you-ready-for-audit.html>
- De Gregorio, G. (2019). Democratising online content moderation: A constitutional framework. *Computer Law & Security Review: The International Journal of Technology Law and Practice*. <https://doi.org/10.1016/j.clsr.2019.105374>
- Deloitte Legal. (2023). Rules for external audit under the RU's Digital Services Act (DSA). *Lexology*. <https://www.lexology.com/library/detail.aspx?g=1301fe0e-46fd-4fea-9444-a68b0bc16901>
- Frosio, G. (2021). Regulatory shift in state intervention: From intermediary liability to responsibility. In Edoardo, C., Amélie, H., & Clara, I (Eds.), *Constitutionalising*

- social media*. Hart Publishing. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3850483
- Kemp, S. (2023, February 9). Digital 2023: Indonesia. *Datareportal*. <https://datareportal.com/reports/digital-2023-indonesia>
- European Commission. (2023). Digital Services Act: Commission designates first set of very large online platforms and search engines. *European Commission*. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_2413
- European Commission. (2023). The Digital Services Act: Ensuring a safe and accountable online environment. *European Commission*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. United Nations Educational, Scientific, and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000233231>
- Legat, O., & Guzy, A. (2023, January 5). Security on the internet. New European initiative - digital services act. *EY*. https://www.ey.com/en_pl/law/security-on-the-internet-new-european-initiative-digital-services-act
- Leong, P. (2022, July 27). Content moderation of social media in Southeast Asia: Contestations and control. *Fulcrum*. <https://fulcrum.sg/content-moderation-of-social-media-in-southeast-asia-contestations-and-control/>
- Palatino, M. (2015, November 29). Will Indonesia's police circular on hate speech suppress freedom of expression?. *Advoc*. <https://advoc.globalvoices.org/2015/11/29/will-indonesias-police-circular-on-hate-speech-suppress-freedom-of-expression/>
- Pirkova, E. (2022, July 6). The Digital Services Act: Your guide to the EU's new content moderation rules. *Access Now*. <https://www.accessnow.org/digital-services-act-eu-content-moderation-rules-guide/>
- PR2Media (2023). *Penyampaian usulan revisi UU ITE kepada Komisi I DPR RI*. [Video]. <https://www.youtube.com/watch?v=bHp5r3wG6xE&t=6s>
- Putri, D. (2021). Apakah semua ujaran kebencian perlu dipidana? Catatan untuk revisi UU ITE. *The Conversation*. <https://theconversation.com/apakah-semua-ujaran-kebencian-perlu-dipidana-catatan-untuk-revisi-uu-ite-156132>
- Santa Clara Principles. (2021). *The Santa Clara principles on transparency and accountability in content moderation*. Electronic Frontier Foundation. <https://santaclaraprinciples.org/open-consultation/>
- Surat Keputusan Bersama Menteri Koordinator Bidang Politik, Hukum, dan Keamanan RI Nomor B-96/HK.00.00/07/2021 tentang Pedoman Implementasi atas Pasal Tertentu dalam UU Informasi dan Transaksi Elektronik (UU ITE)

Surat Keputusan Bersama Menteri Komunikasi dan Informatika RI, Jaksa Agung RI, dan Kepala Kepolisian RI Nomor 229 Tahun 2021 dan Nomor 154 Tahun 2021 tentang Pedoman Implementasi Atas Pasal Tertentu dalam 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

Tremau. (2023, May 31). The DSA & audits. *Tremau*. <https://tremau.com/the-dsa-audits>

United Nations Assembly Resolution 2106. (1965). *International convention on the elimination of all forms of racial discrimination* [General Assembly Resolution 2106 on 21 December 1965]. <https://www.ohchr.org/sites/default/files/cerd.pdf>

United Nations High Commissioner for Human Rights. (2013). *Annual report of the United Nations High Commissioner for Human Rights*. United Nations High Commissioner for Human Rights. https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf

United Nations Human Rights Office of the High Commissioner. (2023). Special rapporteur on freedom of opinion and expression. *United Nations Human Rights Office of the High Commissioner*. <https://www.ohchr.org/en/special-procedures/sr-freedom-of-opinion-and-expression>

US Congress. (2019). Liability for content hosts: An overview of the communication decency act's section 230. *Congressional Research Service*. <https://sgp.fas.org/crs/misc/LSB10306.pdf>

Wendratama, E., Masduki., Rahayu., Suci, P. L., Rianto, P., Aprilia, M. P., Paramastri, M. A., & Adiputra, W. M. (2023). *Pengaturan konten ilegal dan berbahaya di media sosial: Riset pengalaman pengguna dan rekomendasi kebijakan*. Pemantau Regulasi dan Regulator Media. <https://pr2media.or.id/publikasi/pengaturan-konten-ilegal-dan-berbahaya-di-media-sosial-riset-pengalaman-pengguna-dan-rekomendasi-kebijakan/>

Wendratama, E. (2021). Perlindungan data pribadi: Tantangan regulasi di Indonesia. In Masduki (Ed.), *Kebijakan media dan COVID-19 di Indonesia*. Penerbit Komunikasi Universitas Islam Indonesia Program Studi Ilmu Komunikasi. <https://pr2media.or.id/publikasi/kebijakan-media-dan-covid-19-di-indonesia/>

Yanisky-Ravid, S., & Lahav, B. Z. (2017). Public interest vs. private lives – affording public figures privacy, in the digital era: The three principles filtering model. *University of Pennsylvania Journal of Constitutional Law*, 19(5), 1–61. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2931864

Peraturan Perundang-undangan dan Instrumen Internasional

- Undang-Undang No. 29 Tahun 1999 tentang Pengesahan International Convention on The Elimination of All Forms of Racial Discrimination 1965 (Konvensi Internasional tentang Penghapusan Segala Bentuk Diskriminasi Rasial 1965), 1999. https://www.dpr.go.id/dokjdih/document/uu/UU_1999_29.pdf
- Undang-Undang No. 40 Tahun 1999 tentang Pers. <https://peraturan.bpk.go.id/Home/Details/45370/uu-no-40-tahun-1999>
- Undang-Undang No. 12 Tahun 2005 tentang Pengesahan International Covenant on Civil and Political Rights (Kovenan Internasional tentang Hak-Hak Sipil dan Politik), 2005). <https://www.dpr.go.id/doksetjen/dokumen/-Regulasi-UU-No.-12-Tahun-2005-Tentang-Pengesahan-Kovenan-Internasional-Tentang-Hak-Hak-Sipil-dan-Politik-1552380410.pdf>
- Undang-Undang No 19 Tahun 2016 Tentang Perubahan Atas Undang-Undang Nomor 11 Tahun 2008 Tentang Informasi Dan Transaksi Elektronik (2016). <https://web.kominfo.go.id/sites/default/files/users/4761/UU%2019%20Tahun%202016.pdf>
- Peraturan Pemerintah No. 71/2019 tentang Penyelenggaraan Sistem Transaksi Elektronik
- Peraturan Menteri Komunikasi dan Informatika No. 5 Tahun 2020 tentang Penyelenggaraan Sistem Elektronik Lingkup Privat.
- International Convention on the Elimination of All Forms of Racial Discrimination. (1965). International Convention on the Elimination of All Forms of Racial Discrimination. <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-convention-elimination-all-forms-racial>
- International Covenant on Civil and Political Rights (ICCPR). (1966). International Covenant on Civil and Political Rights <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

LAMPIRAN

Usulan Revisi Pasal 15 UU ITE untuk Mendorong Transparansi dan Akuntabilitas Penyelenggara Media Sosial dalam Penanganan Konten Ilegal

(Usulan ini disampaikan PR2Media dalam Rapat Dengar Pendapat Umum yang diadakan Panja Revisi UU ITE Komisi I DPR RI pada 23 Agustus 2023)

No.	Pasal	Bunyi	Usulan dari PR2Media	Penjelasan
1.	Pasal 15	<p>(1) Setiap Penyelenggara Sistem Elektronik harus menyelenggarakan Sistem Elektronik secara andal dan aman serta bertanggung jawab terhadap beroperasinya Sistem Elektronik sebagaimana mestinya.</p> <p>(2) Penyelenggara Sistem Elektronik bertanggung jawab terhadap Penyelenggaraan Sistem Elektroniknya.</p> <p>(3) Ketentuan sebagaimana dimaksud pada ayat (2) tidak berlaku dalam hal dapat dibuktikan terjadinya keadaan memaksa, kesalahan, dan/ atau kelalaian pihak pengguna Sistem Elektronik.</p>	<p>Ditambahkan Ayat (4), yang berbunyi: Dalam melaksanakan ayat (2), penyelenggara media sosial memiliki kewajiban:</p> <p>(a) Membuat tidak dapat diaksesnya informasi elektronik dan/atau dokumen elektronik yang memiliki muatan yang melanggar hukum dengan mengedepankan prinsip kehati-hatian.</p> <p>(b) Menyampaikan secara terbuka tentang tata cara penyelenggara media sosial mengenali dan menentukan informasi elektronik dan/atau dokumen elektronik sebagai informasi elektronik dan/atau dokumen elektronik yang melanggar hukum, baik yang dilakukan oleh sistem otomatis maupun oleh orang yang dipekerjakan penyelenggara media sosial, serta tindakan yang dilakukan oleh penyelenggara media sosial terhadap pelanggaran atau dugaan pelanggaran tersebut.</p> <p>(c) Menyampaikan secara terbuka tentang cara kerja sistem yang digunakan penyelenggara media sosial dalam merekomendasikan informasi elektronik dan/atau dokumen elektronik kepada pengguna dan bagaimana dampak sistem tersebut terhadap informasi elektronik dan/atau dokumen elektronik yang disajikan kepada pengguna.</p> <p>(d) Menyediakan informasi bagi pengguna yang mengalami ketentuan huruf (a) atau penangguhan akun media sosial, yang memuat antara lain informasi elektronik dan/atau dokumen elektronik yang melanggar hukum atau melanggar kebijakan penyelenggara media sosial, yang memuat penjelasan mengenai pelanggaran yang dilakukan, cara penyelenggara media sosial mendeteksi pelanggaran, dan langkah banding yang dapat ditempuh oleh pengguna.</p>	<p>Pengawasan terhadap pelaksanaan ayat (4) dilakukan oleh Pemerintah.</p>

No.	Pasal	Bunyi	Usulan dari PR2Media	Penjelasan
			<p>(e) Menanggapi dan menindaklanjuti aduan dari masyarakat dan pemerintah terkait informasi elektronik dan/atau dokumen elektronik yang memiliki muatan yang melanggar hukum berupa penilaian dan tindakan terhadap objek aduan dengan mengedepankan prinsip kehati-hatian dan keadilan.</p> <p>(f) Mempublikasikan laporan tahunan tentang aduan dari masyarakat dan pemerintah terkait informasi elektronik dan/atau dokumen elektronik yang memiliki muatan yang melanggar hukum serta tindakan pengelola media sosial dalam menindaklanjutinya.</p> <p>(g) Dalam hal ketentuan huruf (a) dilaksanakan, penyelenggara media sosial menyediakan kesempatan banding bagi pengguna, yang memuat antara lain penilaian oleh orang yang tidak terlibat dalam penilaian awal dan/atau kesempatan bagi pengguna untuk memberikan informasi baru.</p> <p>(h) Melaksanakan audit yang diselenggarakan oleh auditor independen setidaknya satu tahun sekali terkait kepatuhan penyelenggara media sosial terhadap ayat (a), ayat (b), ayat (c), ayat (d), ayat (e), atau (f), dan ayat (g) serta menerbitkan laporan audit tersebut secara terbuka.</p> <p>(i) Ketentuan mengenai tata cara kewajiban penyelenggara media sosial sebagaimana dimaksud pada ayat (a), ayat (b), ayat (c), ayat (d), ayat (e), dan ayat (f) diatur dalam peraturan pemerintah.</p>	
2.	Tambahan untuk Bab 1 Ketentuan Umum, Pasal 1	Dalam Undang-Undang ini yang dimaksud dengan: Nomor 1 sampai dengan 23.	Ditambahkan nomor 24. Media sosial adalah sistem elektronik berbasis internet yang memungkinkan para penggunanya saling mengirimkan informasi elektronik dan/atau dokumen elektronik dalam jaringan sistem elektronik yang bersifat terbuka dan diatur oleh penyelenggara media sosial.	—
3.	Penjelasan untuk Pasal 15 ayat (4)	—	Penyelenggara media sosial adalah pihak yang menyelenggarakan sistem elektronik berbasis internet yang memungkinkan para penggunanya saling mengirimkan informasi elektronik dan/atau dokumen elektronik dalam jaringan sistem elektronik yang bersifat terbuka, dengan jumlah pengguna	—

No.	Pasal	Bunyi	Usulan dari PR2Media	Penjelasan
			aktif bulanan di Indonesia di atas 20 juta orang dan/atau media sosial lain yang diusulkan oleh masyarakat karena dianggap penting dan disetujui oleh lembaga berwenang.	
4.	Penjelasan untuk Pasal 15 ayat (4) huruf (h)	—	<p>Pengaturan terkait auditor independen diatur dalam Peraturan Pemerintah, yang memuat antara lain:</p> <ol style="list-style-type: none"> 1. Penyelenggara media sosial memberikan kerjasama dan bantuan yang diperlukan kepada organisasi yang melaksanakan audit, termasuk dengan memberikan akses ke auditor terhadap semua data yang diperlukan untuk melaksanakan audit dengan benar, termasuk, jika relevan, data yang terkait dengan sistem algoritma. 2. Audit harus dilakukan sesuai dengan praktik industri terbaik dan etika profesional yang tinggi serta objektivitas, dengan memperhatikan, standar dan kode praktik audit. 3. Auditor harus menjamin kerahasiaan, keamanan dan keutuhan informasi, seperti rahasia perusahaan, yang diperolehnya dalam melaksanakan tugasnya. 4. Laporan audit harus bisa dibuktikan supaya bisa memberikan penjelasan yang berarti tentang kegiatan yang dilakukan dan kesimpulan yang dicapai. Laporan audit harus membantu menginformasikan, dan jika perlu, memberikan rekomendasi tindakan yang perlu diambil oleh penyelenggara media sosial untuk mematuhi regulasi ini. 5. Laporan audit harus dikirimkan kepada lembaga yang berwenang (Pemerintah). 6. Penyelenggara media sosial harus mengirimkan laporan kepada lembaga yang berwenang (Pemerintah) tentang bagaimana mereka menanggapi laporan audit tersebut. 	<ol style="list-style-type: none"> 1. Audit ini bukanlah audit keuangan, tapi audit ketaatan penyelenggara media sosial terhadap Undang-Undang ini. 2. Audit ini dilaksanakan oleh auditor independen, yaitu organisasi atau perusahaan yang memiliki pengalaman memiliki audit independen.

Catatan:

1. Acuan untuk definisi penyelenggara media sosial: Aichner, T., Grünfelder, M., Maurer, O., & Jegeni, D. (2021). Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. *Cyberpsychology, Behavior and Social Networking*, 24(4), 215-222. Doi: 10.1089/cyber.2020.0134
2. Penggunaan istilah “sistem elektronik yang bersifat terbuka” disesuaikan dengan UU ITE No. 19/2016 Pasal 1 ayat 5 dan 7.
Ayat 5 : Sistem Elektronik adalah serangkaian perangkat dan prosedur elektronik yang berfungsi mempersiapkan, mengumpulkan, mengolah, menganalisis, menyimpan, menampilkan, mengumumkan, mengirimkan, dan/atau menyebarkan Informasi Elektronik.
Ayat 7 : Jaringan Sistem Elektronik adalah terhubungnya dua Sistem Elektronik atau lebih, yang bersifat tertutup ataupun terbuka.
3. Penggunaan istilah “penyelenggara media sosial” disesuaikan dengan UU ITE No. 19/2016 Pasal 1 ayat 6 dan 6a.
Ayat 6 : Penyelenggaraan Sistem Elektronik adalah pemanfaatan Sistem Elektronik oleh penyelenggara negara, Orang, Badan Usaha, dan/atau masyarakat.
Ayat 6a: Penyelenggara Sistem Elektronik adalah setiap Orang, penyelenggara negara, Badan Usaha, dan masyarakat yang menyediakan, mengelola, dan/atau mengoperasikan Sistem Elektronik, baik secara sendiri-sendiri maupun bersama-sama kepada pengguna Sistem Elektronik untuk keperluan dirinya dan/atau keperluan pihak lain.
4. Sejumlah detail dalam ayat (4) huruf (a) hingga (h) mengacu pada Santa Clara Principles <https://santaclaraprinciples.org/open-consultation/> yaitu tata cara yang direkomendasikan oleh organisasi hak asasi, pengacara, dan akademisi kepada platform digital. Sejak kali pertama disuarakan dalam konferensi pengaturan konten di Santa Clara, California pada 2018, prinsip-prinsip ini telah didukung oleh 12 perusahaan digital besar, termasuk Meta, Google, Apple, dan Twitter.
5. Mengenai definisi “penyelenggara media sosial” di Penjelasan untuk Pasal 15 ayat 3, kita tidak bisa menggunakan kuota jumlah pengguna “minimal 10% populasi” seperti di Uni Eropa karena jika demikian maka Twitter tidak bisa masuk ke sini. Itu problematika, jadi muncul angka “20 juta” itu. Di bawah ini data jumlah pengguna media sosial di Indonesia pada awal 2023:
 - a) YouTube : 139 juta
 - b) Facebook : 119 juta
 - c) TikTok : 109,9 juta
 - d) Instagram : 89,1 juta
 - e) Twitter : 24 juta
 - f) LinkedIn : 23 juta

Sumber: <https://datareportal.com/reports/digital-2023-indonesia>

Usulan Pasal untuk Mendorong Transparansi dan Akuntabilitas Penyelenggara Media Sosial dalam Penanganan Konten Ilegal (Sebagai Bagian dari Revisi Total UU ITE)

A. Pengertian dan cakupan media sosial

1. Media sosial adalah sistem elektronik berbasis internet yang memungkinkan para penggunanya saling mengirimkan informasi elektronik dan/atau dokumen elektronik dalam jaringan sistem elektronik yang bersifat terbuka dan diatur oleh penyelenggara media sosial.
2. Penyelenggara media sosial adalah pihak yang menyelenggarakan sistem elektronik berbasis internet yang memungkinkan para penggunanya saling mengirimkan informasi elektronik dan/atau dokumen elektronik dalam jaringan sistem elektronik yang bersifat terbuka, dengan jumlah pengguna aktif bulanan di Indonesia di atas 20 juta orang dan/atau media sosial lain yang diusulkan oleh masyarakat karena dianggap penting dan disetujui oleh lembaga berwenang.

B. Aturan dan kebijakan dari media sosial terkait konten yang dilarang

1. Penyelenggara media sosial menerbitkan aturan dan kebijakan yang jelas dan tepat terkait dengan kapan tindakan akan diambil sehubungan dengan konten atau akun pengguna yang dinilai melanggar aturan perusahaan.
2. Aturan dan kebijakan itu harus membuat pengguna dengan mudah memahami hal-hal berikut:
 - a. Jenis konten apa yang dilarang oleh perusahaan dan akan dihapus, dengan panduan terperinci dan contoh konten yang diizinkan dan tidak diizinkan;
 - b. Jenis konten apa yang akan diambil tindakan oleh perusahaan selain penghapusan, seperti penurunan peringkat algoritmik, dengan panduan terperinci dan contoh pada setiap jenis konten dan tindakan;
 - c. Keadaan ketika perusahaan akan menangguhkan akun pengguna, baik secara permanen maupun sementara.

C. Tata Cara Media Sosial Mengenali dan Menentukan Konten Ilegal

Penyelenggara media sosial memiliki kewajiban:

1. Membuat tidak dapat diaksesnya informasi elektronik dan/atau dokumen

elektronik yang memiliki muatan yang melanggar hukum dengan mengedepankan prinsip kehati-hatian.

2. Menyampaikan secara terbuka tentang tata cara penyelenggara media sosial mengenali dan menentukan informasi elektronik dan/atau dokumen elektronik sebagai informasi elektronik dan/atau dokumen elektronik yang melanggar hukum, baik yang dilakukan oleh sistem otomatis maupun oleh orang yang dipekerjakan penyelenggara media sosial, serta tindakan yang dilakukan oleh penyelenggara media sosial terhadap pelanggaran atau dugaan pelanggaran tersebut.
3. Menyampaikan secara terbuka tentang cara kerja sistem yang digunakan penyelenggara media sosial dalam merekomendasikan informasi elektronik dan/atau dokumen elektronik kepada pengguna dan bagaimana dampak sistem tersebut terhadap informasi elektronik dan/atau dokumen elektronik yang disajikan kepada pengguna.
4. Menyediakan informasi bagi pengguna yang mengalami ketentuan huruf (a) atau penangguhan akun media sosial, yang memuat antara lain informasi elektronik dan/atau dokumen elektronik yang melanggar hukum atau melanggar kebijakan penyelenggara media sosial, yang memuat penjelasan mengenai pelanggaran yang dilakukan, cara penyelenggara media sosial mendeteksi pelanggaran, dan langkah banding yang dapat ditempuh oleh pengguna.
5. Menanggapi dan menindaklanjuti aduan dari masyarakat dan pemerintah terkait informasi elektronik dan/atau dokumen elektronik yang memiliki muatan yang melanggar hukum berupa penilaian dan tindakan terhadap objek aduan dengan mengedepankan prinsip kehati-hatian dan keadilan.
6. Mempublikasikan laporan tahunan tentang aduan dari masyarakat dan pemerintah terkait informasi elektronik dan/atau dokumen elektronik yang memiliki muatan yang melanggar hukum serta tindakan pengelola media sosial dalam menindaklanjutinya.
7. Dalam hal ketentuan huruf (a) dilaksanakan, penyelenggara media sosial menyediakan kesempatan banding bagi pengguna, yang memuat antara lain penilaian oleh orang yang tidak terlibat dalam penilaian awal dan/atau kesempatan bagi pengguna untuk memberikan informasi baru.
8. Melaksanakan audit yang diselenggarakan oleh auditor independen setidaknya satu tahun sekali terkait kepatuhan penyelenggara media sosial terhadap ayat (a), ayat (b), ayat (c), ayat (d), ayat (e), ayat (f), dan ayat (g) serta menerbitkan laporan audit tersebut secara terbuka.

9. Ketentuan mengenai tata cara kewajiban penyelenggara media sosial sebagaimana dimaksud pada ayat (a), ayat (b), ayat (c), ayat (d), ayat (e), dan ayat (f) diatur dalam peraturan pemerintah.

D. Tata Cara Media Sosial Menanggapi Aduan Konten secara Internal

1. Penyelenggara media sosial wajib memiliki sistem penanganan pengaduan internal yang mudah diakses, terstruktur, bersifat transparan, akuntabel, dan menjamin ketepatan waktu penanganan aduan.
2. Sistem penanganan pengaduan internal ini meliputi sistem penerimaan aduan, pemrosesan aduan, penyampaian keputusan, dan pengajuan banding.
3. Penyelenggara media sosial harus mengambil langkah-langkah yang diperlukan untuk mengaktifkan dan memfasilitasi penyampaian aduan yang memuat elemen di bawah ini:
 - a. Penjelasan tentang alasan mengapa informasi tersebut dipermasalahkan sebagai konten ilegal; termasuk aturan hukum yang dilanggar.
 - b. Keberadaan informasi yang diadukan sebagai konten ilegal dengan indikasi yang jelas tentang lokasi elektronik yang tepat dari informasi tersebut, seperti URL, dan, jika perlu, informasi tambahan yang memungkinkan identifikasi konten ilegal yang disesuaikan dengan jenis konten dan jenis layanan hosting tertentu;
 - c. Nama dan alamat email individu atau entitas yang mengirimkan aduan;
 - d. Pernyataan yang menegaskan keyakinan yang dapat dipercaya dari individu atau entitas yang mengirimkan pemberitahuan bahwa informasi dan dugaan yang terkandung di dalamnya adalah akurat dan lengkap.
4. Penanganan aduan tidak diskriminatif, objektif, cermat, adil, proporsional dan menghormati hak pengguna.
5. Penyelenggara media sosial harus memproses setiap pemberitahuan yang mereka terima berdasarkan mekanisme yang jelas dan mengambil keputusan sehubungan dengan pemberitahuan tersebut, secara tepat waktu, cermat, tidak sewenang-wenang dan objektif.
6. Setelah ada pemberitahuan berisi informasi kontak elektronik dari individu atau entitas yang mengirimkannya, penyelenggara media sosial harus, tanpa penundaan yang tidak semestinya, mengirimkan pemberitahuan konfirmasi penerimaan kepada individu atau entitas tersebut.
7. Apabila penyelenggara media sosial menggunakan sarana otomatis untuk pemrosesan atau pengambilan keputusan tersebut, mereka harus memasukkan informasi tentang penggunaan tersebut dalam pemberitahuan.

8. Penanganan aduan dan pembuatan keputusan oleh penyelenggara media sosial wajib dilakukan dan/atau di bawah pengawasan staf yang memenuhi syarat, dan tidak semata-mata atas dasar sarana otomatis.
9. Staf yang menangani aduan dan/atau melakukan fungsi pengaturan konten wajib memahami bahasa suatu wilayah, budaya, dan konteks politik dan sosial dari postingan yang mereka moderasi.
10. Keputusan penyelenggara media sosial terhadap konten ilegal wajib menyertakan penjelasan:
 - a. keputusan apakah akan menghapus atau menonaktifkan akses ke atau membatasi visibilitas informasi atau tidak; atau
 - b. keputusan apakah akan menangguhkan atau mengakhiri penyediaan layanan, seluruhnya atau sebagian, kepada pengguna; atau
 - c. keputusan apakah akan menangguhkan atau menghentikan akun pengguna atau tidak; atau
 - d. keputusan apakah akan menangguhkan, menghentikan, atau membatasi kemampuan untuk memonetisasi informasi yang diberikan oleh pengguna atau tidak.
11. Penyelenggara media sosial harus memberitahu pengadu tanpa penundaan yang tidak semestinya tentang keputusannya terkait pengaduan dan adanya mekanisme banding jika pengadu tidak puas dengan itu.

E. Mekanisme banding

Mekanisme banding internal yang disediakan penyelenggara media sosial mencakup hal-hal sebagai berikut:

1. Proses yang jelas dan mudah diakses oleh pengguna, yang mencakup detail waktu yang harus tertulis secara jelas, supaya pengguna bisa melacak proses kemajuan aduan pengguna.
2. Tinjauan atau penilaian dari orang yang tidak terlibat dalam keputusan awal.
3. Pemahaman bahasa dan konteks budaya yang dimiliki oleh orang yang terlibat dalam proses peninjauan.
4. Kesempatan bagi pengguna untuk menyertakan bukti tambahan untuk mendukung proses banding.
5. Pemberitahuan hasil peninjauan disertai dengan alasan dan pertimbangan yang cukup jelas, agar pengguna memahami keputusan tersebut.

F. Dewan Media Sosial

1. Dalam upaya mengembangkan pengaturan konten di media sosial yang

lebih melayani kepentingan masyarakat, dibentuk Dewan Media Sosial yang independen.

2. Dewan Media Sosial melaksanakan fungsi-fungsi sebagai berikut:
 - a. Memberikan rekomendasi kepada penyelenggara media sosial terkait pengaturan konten.
 - b. Memberikan panduan umum tentang pengaturan konten untuk memastikan praktik tersebut melayani kepentingan masyarakat Indonesia.
 - c. Menjadi lembaga banding administratif bagi pengguna yang tidak puas terhadap keputusan pengaturan konten oleh penyelenggara media sosial.
3. Dewan Media Sosial merupakan lembaga non-struktural (LNS) yang bertanggung jawab kepada Presiden Republik Indonesia.
4. Anggota Dewan Media Sosial terdiri dari:
 - a. Perwakilan pemerintah
 - b. Perwakilan penyelenggara media sosial
 - c. Perwakilan organisasi masyarakat sipil
 - d. Akademisi, tokoh masyarakat, ahli di bidang kebijakan pengaturan media sosial, ahli hukum, dan tokoh agama.
5. Ketua dan Wakil Ketua Dewan Media Sosial dipilih dari dan oleh anggota.
6. Sumber pembiayaan Dewan Media Sosial berasal dari:
 - a. Penyelenggara media sosial
 - b. Bantuan dari negara dan bantuan lain yang tidak mengikat.

G. Laporan tahunan

1. Penyelenggara media sosial harus menyampaikan secara terbuka tentang tata cara dalam mengenali dan menentukan informasi elektronik dan/atau dokumen elektronik yang melanggar hukum, baik yang dilakukan oleh sistem otomatis (*artificial intelligence*) maupun oleh orang yang dipekerjakan (moderator manusia).
2. Penyelenggara media sosial menyampaikan informasi spesifik mengenai jumlah moderator manusia yang dipekerjakan di Indonesia.
3. Penyelenggara media sosial menginformasikan jumlah aduan yang diterima dan tindakan yang dilakukan untuk menangani aduan tersebut.
4. Penyelenggara media sosial menginformasikan tentang konten dan akun yang dikenai sanksi terkait hal-hal berikut:
 - a. Jumlah total konten yang dikenai sanksi dan jumlah total akun ditangguhkan.
 - b. Jumlah banding terhadap keputusan sanksi terhadap konten atau akun.
 - c. Jumlah (atau persentase) banding yang berhasil sehingga konten atau akun dipulihkan, dan jumlah (atau persentase) banding yang gagal;

- d. Jumlah konten atau akun yang dipulihkan oleh penyelenggara media sosial secara proaktif, tanpa banding apa pun, setelah mengetahui bahwa sanksi terhadap konten atau akun tersebut tidak tepat.
 - e. Angka terkait penghapusan dan pembatasan konten yang dilakukan selama periode krisis, seperti selama pandemi COVID-19 dan periode konflik kekerasan.
5. Penyelenggara media sosial menyampaikan laporan mengenai penilaian risiko yang dibawa oleh sistem elektroniknya kepada pengguna dan tindakan mitigasi risiko yang dilakukan oleh penyelenggara media sosial.
 6. Penyelenggara media sosial menyampaikan laporan laporan audit dan pelaksanaan audit yang dilakukan oleh auditor independen.

H. Audit independen terhadap media sosial

1. Untuk mengevaluasi ketaatan penyelenggara media sosial terhadap poin B hingga G di atas, penyelenggara media sosial diaudit oleh organisasi yang melakukan audit atau auditor independen.
2. Penyelenggara media sosial memberikan kerjasama dan bantuan yang diperlukan kepada organisasi yang melaksanakan audit, termasuk dengan memberikan akses ke auditor terhadap semua data yang diperlukan untuk melaksanakan audit dengan benar, termasuk, jika relevan, data yang terkait dengan sistem algoritma.
3. Audit harus dilakukan sesuai dengan praktik industri terbaik dan etika profesional yang tinggi serta objektivitas, dengan memperhatikan, standar dan kode praktik audit.
4. Auditor harus menjamin kerahasiaan, keamanan dan keutuhan informasi, seperti rahasia perusahaan, yang diperolehnya dalam melaksanakan tugasnya.
5. Laporan audit harus bisa dibuktikan supaya bisa memberikan penjelasan yang berarti tentang kegiatan yang dilakukan dan kesimpulan yang dicapai. Laporan audit harus membantu menginformasikan, dan jika perlu, memberikan rekomendasi tindakan yang perlu diambil oleh penyelenggara media sosial untuk mematuhi regulasi ini.
6. Laporan audit harus dikirimkan kepada lembaga yang berwenang (pemerintah dan Dewan Media Sosial).
7. Penyelenggara media sosial harus mengirimkan laporan kepada lembaga yang berwenang (pemerintah dan Dewan Media Sosial) tentang bagaimana mereka menanggapi laporan audit tersebut.
8. Pengaturan terkait auditor independen diatur dalam peraturan teknis.



Pemantau Regulasi dan Regulator Media (PR2Media)
www.pr2media.or.id