

Tata Kelola Moderasi Konten yang Akuntabel dan Demokratis di Indonesia

Analisis terhadap Peraturan Pemerintah (PP) Nomor 71
Tahun 2019 dan Keputusan Menteri Komunikasi dan
Informatika Nomor 522 Tahun 2024





Tata Kelola Konten Moderasi yang Akuntabel dan Demokratis di Indonesia

Analisis terhadap Peraturan Pemerintah (PP) Nomor 71 Tahun 2019 dan Keputusan Menteri Komunikasi dan Informatika Nomor 522 Tahun 2024

Dipublikasikan pada November 2025 oleh Remotivi

Kecuali dinyatakan berbeda, seluruh isi laporan ini dilindungi dengan lisensi Creative Common Attribution 3.0

Ringkasan Eksekutif

Moderasi konten di Indonesia telah berkembang menjadi persoalan politik-hukum yang berdampak langsung pada kebebasan berekspresi dan hak warga untuk memperoleh informasi. Temuan Focus Group Discussion (FGD) mengenai Kepmenkominfo No. 522/2024 menunjukkan bahwa kebijakan pemerintah masih berorientasi pada kontrol alih-alih perlindungan hak. Negara memonopoli penentuan standar moderasi, sementara masyarakat sipil dan pengguna ditempatkan sebagai pihak yang diawasi, bukan sebagai mitra dalam tata kelola ruang digital.

FGD bersama berbagai pemangku kepentingan menyimpulkan perlunya pergeseran menuju kerangka moderasi konten yang lebih sejalan dengan prinsip-prinsip hak asasi manusia. Tiga isu utama mengemuka:

1. Revisi terhadap istilah “meresahkan masyarakat”

Pasal-pasal yang menggunakan istilah seperti “meresahkan masyarakat” dan “mengganggu ketertiban umum” dinilai bertentangan dengan asas kepastian hukum karena tidak memiliki parameter yang terukur. Ketidakjelasan definisi membuka ruang tafsir subjektif aparat dan berpotensi melegitimasi penyensoran terhadap ekspresi kritik atau konten yang mengungkap kekerasan negara. FGD menekankan perlunya menghapus kategori yang moralistik atau, jika tetap dipertahankan, mendefinisikannya melalui konsultasi publik dan indikator yang dapat diverifikasi. Pendekatan berbasis harm principle dianggap lebih selaras dengan komitmen Indonesia terhadap ICCPR.

2. Mendorong model co-regulation dalam moderasi konten

Saat ini seluruh proses evaluasi dan penghapusan konten berada di bawah kendali pemerintah tanpa pengawasan eksternal. Para peserta FGD memandang bahwa struktur ini melemahkan legitimasi dan efektivitas kebijakan. Model co-regulation—melalui dewan pengawas multistakeholder beranggotakan pemerintah, platform, akademisi, dan masyarakat sipil—dinilai mampu memperkuat akuntabilitas, memediasi sengketa, serta meningkatkan kepercayaan publik. Partisipasi masyarakat sipil juga merupakan amanat UU Pembentukan Peraturan Perundang-undangan.



3. Meningkatkan transparansi dan akuntabilitas penghapusan konten

Ketertutupan data mengenai permintaan takedown oleh pemerintah menjadi sumber kecurigaan dan ketegangan publik. Kreator digital, jurnalis, dan pengguna kerap tidak mengetahui alasan penghapusan dan tidak memiliki mekanisme keberatan. FGD merekomendasikan penerapan model transparency reporting sebagaimana Digital Services Act di Uni Eropa, yang mewajibkan pemerintah dan platform mempublikasikan laporan berkala mengenai jumlah, kategori, sumber, dan alasan penghapusan konten. Usulan lain mencakup pengembangan dasbor publik yang menampilkan data real-time tentang operasi moderasi konten.

Kesimpulannya, reformasi moderasi konten di Indonesia harus bergerak dari pendekatan berbasis kontrol menuju tata kelola berbasis hak dan kolaborasi. Revisi konsep “meresahkan masyarakat”, penerapan co-regulation, dan transparansi publik merupakan langkah kunci untuk membangun ekosistem digital yang aman sekaligus melindungi kebebasan berekspresi. Ruang digital yang sehat bukan ditandai oleh absennya perbedaan pendapat, melainkan oleh jaminan bahwa seluruh warga dapat bersuara tanpa rasa takut.


1. Pendahuluan

Demonstrasi yang meluas sejak akhir Agustus 2025 menjadi cerminan nyata bagaimana ruang berekspresi di Indonesia kian menyempit. Aksi yang berawal dari protes atas ketimpangan sosial dan krisis ekonomi itu berujung pada kekerasan aparat, penangkapan sewenang-wenang, serta pembatasan ruang digital. Tim Advokasi untuk Demokrasi (TAUD) mencatat sedikitnya 616 korban, termasuk 399 anak di bawah umur. Sementara Polisi Siber dan Kementerian Komunikasi dan Digital (Komdigi) memblokir 592 akun media sosial yang dianggap memicu provokasi tanpa dasar hukum yang jelas maupun mekanisme banding (SAFENet, 2025a).

Selain itu, berdasarkan laporan Southeast Asia Freedom of Expression Network (SAFENet) terdapat 92 kasus pelanggaran hak digital sejak 25 Agustus 2025 hingga 8 September 2025. Adapun rinciannya, terdapat 19 kasus kriminalisasi terhadap admin-admin media sosial yang kritis dan juga warganet yang bersuara terkait demonstrasi yang dituduh menjadi dalang kerusuhan. Lalu 35 kasus serangan digital berupa *spam call*, *doxing*, *fake order*, dan peretasan. Selain itu, 26 kasus gangguan akses internet berupa pembatasan akses terhadap sejumlah fitur, penangguhan akun, dan pemadaman listrik di beberapa wilayah. Ditambah 16 kasus *over-moderation* yang dimana informasi seputar aksi dan konten kritik lainnya menjadi sasaran. Terakhir, 8 kasus berita bohong dengan dugaan penyebaran informasi yang bertujuan untuk memantik kerusuhan dan 4 kasus ujaran kebencian berbasis rasial dan kekerasan berbasis gender online (SAFENet, 2025b).

Selain berbagai bentuk pelanggaran hak digital di atas, ditemukan juga pembatasan akses terhadap beberapa fitur di media sosial. Misalnya, kesulitan mengunggah ulang *story* yang bernada kritis bahkan hingga aplikasi TikTok mematikan fitur *live streaming*-nya. Berbagai pembatasan dan pelanggaran hak digital yang terjadi dalam konteks aksi demonstrasi Agustus lalu, termasuk praktik moderasi konten, mendapat berbagai kritik dari masyarakat sipil. Mereka menilai praktik moderasi konten yang dijalankan pemerintah dan platform masih cenderung represif serta membuka ruang bagi penyalahgunaan kekuasaan (SAFENet, 2025a). Fenomena ini menunjukkan bahwa kebijakan moderasi konten di Indonesia belum sepenuhnya mampu menyeimbangkan antara kebutuhan menjaga ruang digital yang aman dan perlindungan terhadap hak konstitusional warga untuk berekspresi.

Dalam konteks ini, peran ruang digital semakin penting. Platform digital kini telah menjadi ruang bagi masyarakat untuk berekspresi, berinteraksi, dan membangun wacana publik di Indonesia (Bambang et al., 2025). Tingkat penetrasi internet di Indonesia juga terus meningkat pesat dan kini telah menjangkau sebagian besar




penduduk. Menurut Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), 79,5% masyarakat Indonesia telah terhubung ke internet dan angka ini diproyeksikan naik sekitar 8% pada 2030 (APJII, 2024; Twimbit, 2024). Dari total 285 juta penduduk, lebih dari separuhnya aktif menggunakan media sosial (Kemp, 2025). Perkembangan digital ini memicu semakin banyaknya konten buatan pengguna (*user generated content* atau UGC).

Di satu sisi, hal ini mendorong publik sebagai pengguna untuk mengekspresikan pendapatnya di ruang digital, di sisi lain menghadirkan sejumlah tantangan. Ujaran kebencian, disinformasi, konten yang tidak pantas dan konten berbahaya lainnya yang justru mengancam kondisi demokrasi yang sehat (Humairoh & Nurjannah, 2024). Kondisi di atas akhirnya melahirkan tantangan regulasi untuk menjaga ruang digital tetap aman tanpa mengekang hak warga untuk berbicara dan berpendapat.

Kebebasan berekspresi merupakan hak konstitusional yang dijamin oleh Pasal 28E UUD 1945 dan Pasal 23 Kovenan Internasional Hak Sipil dan Politik (ICCPR), yang telah diratifikasi Indonesia melalui UU No. 12 Tahun 2005. Prinsip ini menegaskan bahwa setiap pembatasan terhadap ekspresi harus didasarkan pada hukum yang jelas, proporsional, dan diperlukan untuk melindungi kepentingan publik yang sah.

Di sisi lain, pemerintah Indonesia melalui Kementerian Komunikasi dan Digital (Kemkomdigi) telah berupaya mengatur ruang digital dengan menempatkan UU Informasi dan Transaksi Elektronik (ITE) serta berbagai turunannya sebagai dasar hukum utama. Regulasi ini, termasuk Permenkominfo No. 5/2020 tentang Penyelenggara Sistem Elektronik Lingkup Privat, mewajibkan seluruh platform digital baik lokal maupun asing untuk mendaftar, mematuhi permintaan penghapusan konten, dan menyerahkan data pengguna bagi kepentingan penegakan hukum. Aturan tersebut memberi Kemkomdigi kewenangan besar untuk meminta penghapusan konten yang dianggap melanggar hukum nasional, seperti ujaran kebencian atau hasutan publik, dengan tenggat penanganan empat jam untuk kasus mendesak dan dua puluh empat jam untuk kasus biasa. Kewajiban ini diperkuat melalui Kepmenkominfo No. 522/2024, yang memperkenalkan *Sistem Kepatuhan Moderasi Konten (SAMAN)* sebagai mekanisme pemantauan terpusat terhadap kepatuhan platform.

Meskipun bertujuan menciptakan ruang digital yang aman, pendekatan ini kerap menuai kritik dari masyarakat. Terdapat sejumlah kekhawatiran dari berbagai kalangan terkait regulasi yang sudah ada saat ini. Rezim regulasi moderasi konten yang menekankan penyensoran dan pembatasan, belum terjaminnya keamanan data dari sistem SAMAN, regulator yang tersentralisasi pada pihak pemerintah, minimnya transparansi, absennya mekanisme banding bagi publik, hingga adanya



potensi mengancam kebebasan pers (AJI, 2025; CIPS, 2025a). Selain itu, salah satu hal krusial lainnya yang kerap dikritik adalah ambiguitas dari definisi konten yang dianggap melanggar. Pasal 9 ayat 4 dalam Permenkominfo No. 5/2020 berbunyi:

Informasi Elektronik dan/atau Dokumen Elektronik yang dilarang sebagaimana dimaksud pada ayat (3) dengan klasifikasi:

- a. melanggar ketentuan peraturan perundang-undangan;*
- b. meresahkan masyarakat dan mengganggu ketertiban umum; dan*
- c. memberitahukan cara atau menyediakan akses terhadap Informasi Elektronik dan/atau Dokumen Elektronik yang dilarang.*

Tidak jelasnya peraturan perundang-undangan yang termasuk dalam poin a serta batasan dari poin b mengenai konten “meresahkan masyarakat dan mengganggu ketertiban umum” berpotensi membuka peluang kriminalisasi berlebihan dan membatasi kebebasan berekspresi masyarakat.

Berangkat dari berbagai persoalan di atas, kebijakan moderasi konten di Indonesia perlu ditinjau ulang agar berpijak pada perlindungan hak konstitusional warga untuk berekspresi. Sejumlah negara demokratis telah mengembangkan model tata kelola yang lebih seimbang dengan menempatkan prinsip transparansi, akuntabilitas, dan hak banding sebagai fondasi utama dalam setiap proses moderasi konten (Toptchiyska, 2023; Galli et al., 2022). Sehingga policy brief ini pun bertujuan untuk memberikan berbagai perspektif komparatif dengan berbagai negara demokratis. Selain itu, policy brief ini pun bertujuan untuk mengakomodasi suara berbagai stakeholder, seperti pelaku industri, akademisi, hingga masyarakat sipil, untuk merumuskan kerangka regulasi moderasi konten yang berkeadilan dan sejalan dengan prinsip-prinsip demokratis. Tujuan akhirnya adalah mendorong terbentuknya kerangka regulasi yang menjamin keamanan ruang digital tanpa mengorbankan kebebasan berekspresi, memperkuat mekanisme pengawasan yang independen, serta memastikan setiap warga memiliki hak atas ruang digital yang inklusif, terbuka, dan akuntabel.



2. Metodologi

Penelitian ini dilakukan melalui pendekatan kualitatif dengan metode studi pustaka dan diskusi terarah (*focused group discussion/FGD*). Studi pustaka dilakukan untuk mengkaji literatur yang relevan terkait isu moderasi konten, hak digital, tata kelola digital, dan kebebasan pers. Selain itu, diskusi terarah dilaksanakan secara tatap muka pada bulan September 2025, melibatkan para pemangku kepentingan dari berbagai latar belakang, termasuk pengamat, organisasi jurnalis, kreator digital, hingga perwakilan lembaga yang fokus pada berbagai isu, seperti bidang hak digital, kebijakan publik, serta reformasi hukum. Diskusi ini bertujuan untuk memperoleh perspektif multidisipliner dalam merumuskan temuan dan rekomendasi kebijakan moderasi konten di Indonesia.

3. Ekosistem Moderasi Konten di Indonesia dan Masalahnya

Dalam satu dekade terakhir, kebijakan negara terhadap ruang digital di Indonesia menunjukkan kecenderungan yang semakin intervensif. Pemerintah, melalui Kementerian Komunikasi dan Digital (Komdigi), tidak lagi sekadar berperan sebagai penyedia infrastruktur dan fasilitator transformasi digital, melainkan juga sebagai penentu batas-batas ekspresi di ruang maya. Pergeseran fungsi ini tampak jelas dalam serangkaian regulasi yang membentuk ekosistem moderasi konten nasional, mulai dari Undang-Undang Informasi dan Transaksi Elektronik (UU ITE) hingga Keputusan Menteri Komunikasi dan Informatika (Kepmenkominfo) No. 522 Tahun 2024.

Kebijakan ini pada mulanya dirancang untuk menanggulangi konten digital yang berpotensi membahayakan publik, seperti terorisme, pornografi anak, dan perjudian daring. Namun dalam praktiknya, cakupan tersebut meluas hingga menyentuh wilayah ekspresi yang sah secara hukum, seperti kritik terhadap pemerintah atau liputan jurnalistik. Akibatnya, batas antara moderasi dan sensor menjadi semakin kabur, sementara tanggung jawab terhadap kebebasan berekspresi bergeser dari ruang publik ke ruang administratif negara.

3.1 Arsitektur Regulasi Moderasi Konten

Kerangka regulasi moderasi konten di Indonesia dibangun secara bertahap melalui sejumlah peraturan yang saling berkaitan dan memperkuat peran negara dalam mengawasi lalu lintas informasi digital.

Fondasi hukumnya adalah Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE), yang kemudian direvisi pada 2016 dan 2024. UU ini menjadi tonggak pertama bagi keterlibatan negara dalam mengatur komunikasi di ruang digital. Beberapa pasalnya mengatur larangan terhadap muatan yang dianggap melanggar kesusilaan, perjudian, penghinaan, penyebaran kebencian berbasis SARA, hingga penyebaran berita bohong. Walaupun dimaksudkan untuk melindungi masyarakat dari konten berbahaya, UU ITE sejak awal menimbulkan kekhawatiran karena definisi “konten ilegal” di dalamnya sangat kabur dan telah digunakan dalam berbagai kasus untuk menjerat ekspresi warga yang sah secara hukum.

Untuk menurunkan amanat UU ITE, pemerintah menerbitkan Peraturan Pemerintah (PP) No. 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik (PSTE). Regulasi ini memperluas mandat pengawasan negara terhadap penyelenggara sistem elektronik (PSE)—baik perusahaan lokal maupun asing—dan memberi kewenangan kepada pemerintah untuk memutus akses terhadap informasi elektronik yang mengandung muatan terlarang. Namun PP ini tidak memberikan batasan jelas mengenai apa yang dimaksud dengan “muatan terlarang” atau bagaimana prosedur penilaiannya dilakukan. Kekosongan definisi ini membuka ruang interpretasi yang luas dan menjadikan aparat administratif sebagai penentu sah tidaknya suatu ekspresi digital.

Kekosongan tersebut kemudian diisi oleh Peraturan Menteri Komunikasi dan Informatika (Permenkominfo) No. 5 Tahun 2020 tentang Penyelenggara Sistem Elektronik Lingkup Privat, yang secara eksplisit membentuk mekanisme moderasi konten nasional. Peraturan ini mewajibkan semua platform digital untuk mendaftar ke pemerintah dan menindaklanjuti permintaan penghapusan konten yang dianggap melanggar hukum Indonesia. Dalam kasus yang dinilai mendesak, platform harus menurunkan konten dalam waktu empat jam; untuk kasus lain, dalam 24 jam. Kegagalan mematuhi aturan ini dapat berujung pada sanksi administratif, denda, atau bahkan pemblokiran layanan.

Untuk memperkuat pelaksanaan kebijakan ini, pemerintah menerbitkan Kepmenkominfo No. 172 Tahun 2024, yang kemudian direvisi menjadi Kepmenkominfo No. 522 Tahun 2024. Regulasi ini memperkenalkan *Sistem Kepatuhan Moderasi Konten* atau SAMAN, sebuah portal digital yang mengintegrasikan proses pelaporan, verifikasi, dan penghapusan konten. Melalui sistem ini, pemerintah dapat memantau dan menegur platform secara langsung,



serta menjatuhkan sanksi apabila mereka tidak mematuhi perintah penghapusan konten.


Secara formal, SAMAN diklaim sebagai bentuk modernisasi tata kelola konten yang lebih efisien dan transparan. Namun, sebagaimana dicatat dalam kajian CIPS (2025b), desainnya justru memperkuat sentralisasi kekuasaan di tangan pemerintah. Prosedur banding hanya dapat dilakukan setelah konten dihapus, dan lembaga yang memproses banding adalah institusi yang sama yang mengeluarkan perintah penghapusan. Dengan demikian, sistem ini meniadakan prinsip *due process of law* yang menjadi dasar perlindungan hak asasi manusia. Selain itu, kategori pelanggaran seperti “konten yang meresahkan masyarakat” atau “mengganggu ketertiban umum” tetap dipertahankan tanpa definisi yang jelas, sehingga membuka ruang subjektivitas dan penyalahgunaan tafsir oleh aparat.

3.2 Masalah Struktural dan Implikasi terhadap Hak Asasi Manusia

Kelemahan utama dari keseluruhan kerangka regulasi ini terletak pada ketimpangan kekuasaan antara negara, platform, dan warga negara. Regulasi yang ada memberi kewenangan besar bagi pemerintah untuk menentukan batas ekspresi publik, tetapi hampir tidak menyediakan mekanisme pengawasan independen terhadap penggunaan kewenangan tersebut. Akibatnya, negara menjadi hakim sekaligus juri dalam menentukan apakah sebuah konten layak beredar di ruang digital.

Permenkominfo No. 5/2020 dan Kepmen 522/2024 juga menempatkan beban besar pada penyelenggara sistem elektronik. Tenggat waktu empat jam untuk menurunkan konten yang dianggap mendesak mendorong banyak platform untuk mengambil jalan aman dengan melakukan *over-moderation*, yakni menghapus konten yang ambigu sebelum sempat diverifikasi kebenarannya. Dalam konteks ini, kepatuhan administratif menjadi lebih penting daripada perlindungan terhadap kebebasan berekspresi.

Selain masalah substansial, aspek teknis dari sistem SAMAN juga menimbulkan kekhawatiran serius. Hingga akhir 2024, sistem ini belum memiliki sertifikasi keamanan data internasional, sementara rekam jejak pemerintah dalam pengelolaan data publik masih lemah. Sekitar 70% kasus kebocoran data di Indonesia pada 2023 justru terjadi di lembaga pemerintah. Jika sistem moderasi konten mengharuskan platform untuk menyerahkan data pengguna yang terlibat dalam laporan, maka potensi pelanggaran privasi menjadi semakin besar.



Hasil *Focus Group Discussion (FGD)* tentang Kepmen 522 memberikan gambaran konkret mengenai bagaimana kebijakan moderasi konten ini diimplementasikan di lapangan dan dampaknya terhadap berbagai sektor.


Dalam diskusi tersebut, direktur perempuan dari lembaga yang fokus di bidang hak digital menuturkan pengalaman ketika fitur *live* TikTok tiba-tiba tidak dapat digunakan selama demonstrasi besar pada Agustus 2025. Ia melihat situasi itu sebagai bentuk pembatasan hak publik untuk mengakses informasi secara langsung. “Orang jadi kesulitan melihat situasi lapangan secara langsung,” ujarnya. Baginya, hilangnya fitur tersebut memperlihatkan bagaimana regulasi moderasi konten dapat dengan mudah digunakan untuk menekan arus informasi di momen-momen politik yang sensitif (FGD Kepmen 522, 25/09/2025).

Dampak serupa juga dirasakan oleh kalangan jurnalis. Salah satu pegiat asosiasi jurnalis menceritakan bahwa sejumlah akun media, misalnya *Pikiran Rakyat*, tiba-tiba kehilangan akses fitur siaran langsung tak lama setelah Komdigi memanggil pihak TikTok pada 29 Agustus 2025. “Akun kami di-*banned*, fitur *live* hilang tanpa pemberitahuan,” ujarnya. Ia menilai bahwa kebijakan tersebut bukan hanya membatasi kerja jurnalistik, tetapi juga menciptakan suasana takut bagi jurnalis yang meliput isu-isu kritis (FGD Kepmen 522, 25/09/2025).

Tidak hanya jurnalis, kreator digital, mengalami langsung pelaporan konten oleh Komdigi. Ia menerima notifikasi dari platform X (Twitter) bahwa unggahan media mereka tentang kekerasan politik 1998 dilaporkan pemerintah sebagai pelanggaran hukum. “Dari situ kami sadar bahwa pengawasan bisa terjadi kapan saja,” ujar salah satu kreator digital di bidang edukasi sejarah. Dirinya menilai praktik seperti ini memperlihatkan pergeseran fungsi moderasi konten menjadi alat pengawasan politik, di mana konten kritis berisiko dihapus atas dasar kepatuhan administratif (FGD Kepmen 522, 25/09/2025).

Sementara itu, salah satu manajer perempuan dari lembaga advokasi reformasi hukum mengkritik cacat formil dari Kepmen 522. Ia menegaskan bahwa bentuk hukum *Keputusan Menteri* seharusnya tidak digunakan untuk mengatur norma umum. “Pembatasan hak warga seperti kebebasan berekspresi harus diatur dalam Peraturan Pemerintah, bukan dalam Keputusan Menteri,” ujarnya. Kesalahan ini, menurutnya, membuat publik kehilangan dasar untuk mengajukan *judicial review* karena Kepmen tidak dapat diuji di Mahkamah Agung (FGD Kepmen 522, 25/09/2025).

Menariknya, salah satu pengamat kebebasan digital di Indonesia menempatkan persoalan ini dalam konteks historis yang lebih luas. Ia berpendapat bahwa kebijakan moderasi konten di Indonesia selalu “berjalan dulu, baru dibuatkan regulasinya.” Ia mencontohkan sistem *Trust Positif* yang pernah dijalankan sebelum memiliki dasar hukum yang kuat, dan kini pola itu berulang dalam



implementasi SAMAN. Dirinya menambahkan bahwa bentuk penyensoran kini tidak hanya terjadi di tingkat konten, tetapi juga pada lapisan infrastruktur, seperti pemblokiran domain melalui DNS Nusantara yang dikelola oleh APJII (FGD Kepmen 522, 25/09/2025). “Ini bentuk sensor yang lebih tersembunyi,” ujarnya, “karena dilakukan atas nama keamanan jaringan, padahal efeknya sama: membatasi akses publik terhadap informasi.”

Dari keseluruhan pandangan dalam FGD, tampak bahwa kebijakan moderasi konten di Indonesia menciptakan efek berantai yang kompleks: membatasi akses informasi publik, menekan kerja jurnalis, sekaligus melemahkan akuntabilitas hukum. Di balik narasi formal tentang keamanan dan ketertiban digital, terdapat praktik administratif yang berpotensi mempersempit ruang demokrasi.

Serangkaian fakta-fakta di atas memberi kita indikasi kuat bahwa ekosistem moderasi konten di Indonesia dibangun di atas kerangka regulasi yang terpusat, tidak transparan, dan minim akuntabilitas publik. Dalam konteks hak asasi manusia, kondisi ini mengancam tiga prinsip mendasar: kebebasan berekspresi, hak atas informasi, dan hak atas perlindungan hukum yang adil. Kebijakan yang lahir dengan dalih keamanan digital justru berpotensi mengikis kepercayaan publik terhadap negara dan melemahkan fondasi demokrasi digital.

Pemerintah perlu mengubah pendekatan dari model berbasis penyensoran menuju model berbasis perlindungan hak. Hal ini mencakup kejelasan definisi hukum, transparansi proses, mekanisme banding independen, serta pelibatan multipihak mulai dari masyarakat sipil, industri, maupun komunitas pengguna. Inilah prasyarat yang dibutuhkan agar moderasi konten tidak berubah menjadi instrumen pembatasan demokrasi digital yang justru menutup ruang partisipasi publik yang hendak dijaganya.

4. Praktik Moderasi Konten di Negara-Negara Demokratis

Perkembangan pesat ruang digital telah mengubah cara negara-negara demokratis merespons persoalan kebebasan berekspresi dan keamanan daring (Galli et al., 2022). Ledakan konten buatan pengguna menghadirkan dilema klasik: bagaimana menekan penyebaran konten berbahaya tanpa mengekang hak publik untuk berbicara. Di tengah situasi tersebut, banyak negara demokratis merumuskan kerangka regulasi yang berusaha menyeimbangkan antara perlindungan terhadap individu dan penghormatan terhadap kebebasan sipil. Pengalaman mereka penting untuk dikaji, terutama karena Indonesia saat ini masih mengandalkan pendekatan yang bersifat administratif dan terpusat. Melalui pembelajaran dari praktik di Uni Eropa, Amerika Serikat, Inggris, dan Kanada, bagian ini berupaya menelusuri bagaimana negara-negara demokratis membangun sistem moderasi



konten yang lebih akuntabel, transparan, dan sejalan dengan prinsip-prinsip demokratis.

4.1 Konsep Konten Ilegal & Konten Berbahaya

Negara-negara demokratis membedakan secara tegas antara konten ilegal (*illegal content*) dan konten berbahaya (*harmful content*). Kategori pertama merujuk pada konten yang secara eksplisit melanggar hukum pidana, seperti terorisme, pornografi anak, ujaran kebencian yang menghasut kekerasan, dan pelanggaran hak cipta. Sementara kategori kedua mencakup konten yang dinilai berpotensi merugikan masyarakat tetapi tidak secara langsung melanggar hukum, misalnya misinformasi, pelecehan daring, atau konten ekstrem yang menimbulkan bahaya sosial.

Di Amerika Serikat, konsep kebebasan berekspresi dilindungi kuat oleh First Amendment Konstitusi AS, yang menegaskan bahwa negara tidak dapat membuat aturan yang mengekang kebebasan berekspresi (United States, n.d.). Meskipun begitu, ada beberapa kategori konten ilegal yang turut diatur dalam hukum pidana federal, seperti eksploitasi seksual anak, konten cabul (*obscene or harassing*) atau konten *harassment* yang tidak memiliki *consent* orang dewasa dan melibatkan anak-anak, penipuan, ujaran kebencian yang berujung kekerasan, serta dukungan terhadap organisasi teroris. Adapun untuk konten-konten yang berbahaya (*harmful content*) di luar dari yang diatur hukum pidana, platform diberikan kewenangan untuk melakukan moderasi konten secara sukarela. Ketentuan tersebut diatur dalam Section 230 of the Communications Decency Act (47 U.S.C. §230). Sehingga, negara hanya dapat bertindak jika ada pelanggaran hukum yang jelas, sementara tanggung jawab etis dan teknis untuk menekan dampak konten berbahaya diserahkan kepada platform melalui *community standards* dan *transparency reporting*.

Berbeda dengan Amerika Serikat, Inggris melalui Online Safety Act (OSA) 2023 menerapkan pembedaan yang lebih regulatif antara *illegal content* dan *harmful content*. Berdasarkan pasal 59 OSA, konten ilegal meliputi materi yang merupakan tindak pidana relevan (*relevant offence*) dan tindak pidana prioritas (*priority offence*) (United Kingdom, 2024). Adapun tindak pidana prioritas ini terbagi menjadi tiga, yakni konten eksploitasi dan pelecehan seksual anak, terorisme, dan tindak pidana prioritas lainnya, seperti pelecehan seksual, ancaman bunuh diri, menguntit, perilaku mengancam, tindak pidana yang diperburuk oleh agama dan ras, tindak pidana terkait penjualan senjata, dan tindak pidana lainnya. Sementara *harmful content* hanya konten-konten yang berbahaya bagi anak (*content harmful to children*), seperti konten yang abusif, konten yang mendorong, mempromosikan, atau memberikan instruksi untuk tindakan kekerasan serius terhadap seseorang




dan aksi berbahaya, konten yang mendorong penggunaan zat yang berbahaya, dan lain sebagainya.

Sehingga secara umum, OSA membagi konten melanggar menjadi dua, yakni konten ilegal dan konten berbahaya bagi anak. Seluruh konten ilegal didefinisikan berdasarkan ketentuan hukum pidana yang telah ada sebelumnya, sedangkan kategori konten berbahaya dibangun atas prinsip perlindungan anak (United Kingdom, 2023). Sehingga logika dasar yang mendasari model ini bukanlah pembatasan ekspresi, melainkan perlindungan hak, sehingga memperkecil risiko *over-moderation* dan menjaga keseimbangan antara keamanan daring dan kebebasan berekspresi.

Kanada juga mengatur perbedaan antara konten ilegal dan berbahaya melalui Online Harms Act (Bill C-63). Dalam aturan tersebut, konten berbahaya meliputi: (a) konten intim yang disebarakan tanpa persetujuan, (b) konten yang mengeksploitasi anak, (c) konten yang mendorong anak untuk menyakiti diri sendiri, (d) konten yang digunakan untuk mengintimidasi anak, (e) konten yang menghasut kebencian, (f) konten yang menghasut kekerasan, dan (g) konten yang menghasut ekstremisme kekerasan atau terorisme (Canada, 2024). Sementara itu, *illegal content* tetap diatur dalam Criminal Code dan Canadian Human Rights Act. Senada dengan di Inggris, logika dasar dari klasifikasi ini berpijak pada perlindungan terhadap hak dan keselamatan individu, bukan pada pembatasan ekspresi. Semua kategori *illegal content* didasarkan pada hukum pidana yang telah ada sebelumnya, sedangkan *harmful content* dirumuskan dalam kerangka perlindungan kelompok rentan, khususnya anak dan korban kekerasan daring. Pendekatan ini menegaskan bahwa kebijakan moderasi konten harus bersifat *rights-based*, di mana negara dan platform memiliki tanggung jawab untuk memastikan ruang digital yang aman tanpa membungkam kebebasan berekspresi.

Sementara itu, Uni Eropa melalui Digital Services Act (Regulation (EU) 2022/2065) menegaskan prinsip *illegal content must go, harmful content must flow responsibly*. Artinya, platform wajib menindak konten ilegal berdasarkan kerangka hukum nasional, tetapi tidak boleh secara sewenang-wenang membatasi konten yang tidak melanggar hukum. Untuk konten berbahaya seperti disinformasi, DSA memperkenalkan mekanisme *code of conduct* sukarela dan *risk mitigation obligation* bagi *Very Large Online Platforms (VLOPs)* agar menilai dampak sistemik konten terhadap keamanan publik dan kebebasan berekspresi (European Union, 2022). Pendekatan ini mencerminkan filosofi khas Eropa yang menyeimbangkan *fundamental rights* dengan *public order*, yang menempatkan transparansi dan audit independen sebagai pengawas utama.

Dari keempat yurisdiksi tersebut, terlihat bahwa logika yang dibangun selalu berakar pada keseimbangan antara perlindungan publik dan hak konstitusional




untuk berekspresi. Amerika Serikat menekankan pembatasan minimum terhadap ekspresi; Inggris dan Kanada memusatkan perhatian pada pencegahan risiko dan perlindungan kelompok rentan; sedangkan Uni Eropa berupaya menata arsitektur tanggung jawab digital dengan memperkuat kewajiban transparansi. Model-model ini menunjukkan bahwa penanganan konten berbahaya tidak harus berarti pembatasan ekspresi, tetapi penguatan tata kelola yang berbasis akuntabilitas, proporsionalitas, dan hak pengguna.

4.2 Bagaimana Negara-Negara Demokratis Mengatur Konten Berbahaya dan Ilegal

Perbedaan utama antara konten ilegal dan konten berbahaya tercermin jelas dalam cara negara-negara demokratis membangun mekanisme regulasi dan kelembagaan. Negara-negara tersebut berupaya memastikan bahwa setiap intervensi terhadap ekspresi publik memiliki dasar hukum yang tegas, dilakukan oleh lembaga independen, dan disertai mekanisme akuntabilitas publik yang kuat. Pendekatan ini menempatkan perlindungan terhadap kebebasan berekspresi bukan sebagai hambatan, tetapi sebagai prasyarat bagi tata kelola digital yang demokratis.

Di Amerika Serikat, moderasi konten beroperasi dalam kerangka perlindungan kuat terhadap kebebasan berekspresi sebagaimana dijamin oleh *First Amendment*. Negara tidak memiliki kewenangan untuk menentukan atau memaksa penghapusan konten, kecuali terhadap ekspresi yang secara eksplisit dikategorikan sebagai pelanggaran hukum pidana. Melalui *Section 230 of the Communications Decency Act (47 U.S.C. §230)*, penyedia layanan daring memperoleh kekebalan hukum atas konten yang diunggah pengguna, sekaligus kebebasan untuk melakukan moderasi sukarela tanpa kehilangan perlindungan hukum tersebut. Dalam kasus konten yang melanggar hukum pidana, penyedia layanan wajib melaporkan kepada National Center for Missing and Exploited Children (NCMEC) untuk kasus eksploitasi anak, atau kepada Department of Justice (DOJ) dan pengadilan terkait untuk pelanggaran lain. Platform juga diwajibkan menyimpan bukti digital hingga satu tahun untuk keperluan penyelidikan. Penegakan hukum dilakukan oleh lembaga negara melalui proses pidana, sementara mekanisme banding atas penghapusan konten berada di ranah kebijakan internal platform atau melalui peradilan umum. Pola ini menunjukkan bahwa model moderasi konten di Amerika Serikat bersifat *self-regulatory* dengan batas hukum yang jelas, di mana negara hanya turun tangan ketika terjadi pelanggaran hukum pidana yang konkret.


Berbeda dengan Amerika Serikat yang menekankan *self-regulation*, Inggris menerapkan model *co-regulation* melalui Online Safety Act (OSA) 2023, yang



menempatkan tanggung jawab utama pada platform sekaligus memberi wewenang pengawasan kepada lembaga independen, yakni Ofcom. OSA mewajibkan setiap penyedia layanan daring—baik *user-to-user services* maupun *search services*—untuk memiliki *duties of care*, termasuk penilaian risiko (*risk assessment*) terhadap konten ilegal dan konten berbahaya bagi anak. Hasil penilaian ini menjadi dasar bagi platform untuk menyusun sistem moderasi internal, termasuk mekanisme pelaporan pengguna (*user reporting*), penghapusan konten (*take-down*), dan prosedur banding bagi pengguna yang kontennya dihapus. Ketika sebuah konten terdeteksi melanggar, platform wajib menanganinya sesuai hasil *risk assessment* dan pedoman yang tercantum dalam Code of Practice yang disusun oleh Ofcom (United Kingdom, 2024).

Jika platform gagal menindaklanjuti pelaporan atau penilaian risiko secara memadai, Ofcom dapat melakukan proses penegakan bertahap (*step-by-step enforcement*). Tahap pertama berupa pemberitahuan resmi (*notice*) untuk memperbaiki pelanggaran. Bila diabaikan, Ofcom mengeluarkan keputusan konfirmasi (*confirmation decision*) yang bersifat mengikat. Pelanggaran lanjutan dapat dikenai sanksi administratif berupa denda hingga 10% dari pendapatan global, dan dalam kasus berat, pemutusan akses terhadap layanan di wilayah Inggris (United Kingdom, 2023). Pendekatan ini menunjukkan bahwa logika utama moderasi konten di Inggris bukanlah pembatasan ekspresi, tetapi perlindungan hak dan keselamatan pengguna melalui mekanisme yang proporsional dan berbasis risiko (*risk-based regulation*).

Di Kanada, tata kelola moderasi konten diatur melalui *Online Harms Act* (Bill C-63, 2024) yang memperkenalkan pendekatan ko-regulatif berbasis tanggung jawab berlapis antara negara, platform, dan regulator independen. Pemerintah federal menetapkan kerangka umum perlindungan melalui undang-undang ini sekaligus membentuk lembaga pengawas bernama Digital Safety Commission yang berfungsi sebagai otoritas penegak dan penafsir kebijakan. Platform digital diwajibkan menyusun Rencana Keamanan Digital (*Digital Safety Plan*) yang berisi kebijakan internal, prosedur pelaporan, dan langkah mitigasi risiko sesuai standar yang ditetapkan oleh Komisi. Ketika sebuah konten dilaporkan melanggar kategori *harmful content*, laporan tersebut dapat diajukan baik ke operator platform maupun langsung ke Komisi. Platform kemudian melakukan penonaktifan sementara (*temporary inaccessibility*) terhadap konten yang dilaporkan sambil melakukan peninjauan. Dalam tahap ini, baik pelapor maupun pengunggah diberikan kesempatan untuk menyampaikan representasi atau pembelaan. Jika setelah peninjauan awal ditemukan alasan yang cukup, konten tersebut akan tetap dinonaktifkan; sebaliknya, jika tidak, konten akan dikembalikan (*reinstated*). Pengguna dapat mengajukan banding terhadap keputusan platform ataupun Komisi. Apabila sebuah platform berulang kali mengabaikan keputusan Komisi atau tidak mematuhi standar moderasi yang diwajibkan, lembaga tersebut



berwenang mengenakan denda administratif yang signifikan (Canada, 2024). Melalui mekanisme ini, Kanada mencoba menyeimbangkan antara perlindungan terhadap pengguna dari konten berbahaya dan penegakan hak kebebasan berekspresi melalui prosedur yang transparan dan berlapis pengawasan.

Uni Eropa mengatur tata kelola moderasi konten melalui Digital Services Act (DSA), Regulation (EU) 2022/2065, yang menekankan prinsip transparansi, akuntabilitas, dan perlindungan hak pengguna dalam proses penghapusan konten. DSA menetapkan standar umum yang berlaku bagi seluruh penyedia layanan digital di wilayah Uni Eropa dan mewajibkan setiap negara anggota menunjuk Digital Services Coordinator (DSC) sebagai otoritas nasional pengawas. Platform diwajibkan membangun mekanisme *notice-and-action* yang memungkinkan pengguna maupun *trusted flagger* melaporkan konten ilegal atau berbahaya. Setelah menerima laporan, platform dapat menonaktifkan atau menghapus konten sementara disertai pemberitahuan kepada pengguna yang terdampak. Pengunggah diberikan hak banding baik secara internal di platform maupun ke lembaga independen. Dalam kasus serius, DSC dapat mengeluarkan perintah penghapusan (*removal order*), dan untuk kasus lintas negara atau yang melibatkan *very large online platforms (VLOPs)*, kewenangan beralih ke Komisi Eropa. DSA juga mengatur sanksi berupa denda administratif bagi platform yang secara berulang melanggar kewajiban transparansi atau gagal menindak laporan yang sah (European Union, 2022). Dengan model ini, Uni Eropa menyeimbangkan antara intervensi negara dan tanggung jawab platform melalui mekanisme pengawasan berlapis serta prosedur banding yang menjamin perlindungan kebebasan berekspresi.


Secara keseluruhan, model moderasi konten di berbagai negara demokratis memperlihatkan pola yang sama: negara tidak secara langsung mengendalikan isi konten, tetapi mengatur mekanisme tata kelola yang menuntut platform bertanggung jawab terhadap risiko konten berbahaya sekaligus menjaga hak warga untuk berekspresi. Amerika Serikat menekankan kebebasan berekspresi dan perlindungan hukum bagi platform; Inggris dan Uni Eropa menggunakan pendekatan ko-regulatif dengan prinsip *duties of care* dan *notice-and-action*; sementara Kanada mengembangkan model yang lebih partisipatif dengan fokus pada perlindungan korban dan pengawasan independen. Meski tiap negara memiliki logika regulatif berbeda, keseluruhannya menunjukkan upaya untuk menyeimbangkan keamanan digital dan hak asasi manusia, dua nilai yang menjadi fondasi utama demokrasi digital modern.

4.3 Bagaimana Negara-Negara Demokratis Menjamin Akuntabilitas dalam Pengaturan Moderasi Konten

Prinsip akuntabilitas menjadi salah satu pilar utama dalam tata kelola moderasi konten di negara-negara demokratis. Meskipun pendekatannya berbeda, keseluruhannya berpijak pada satu gagasan bahwa kekuasaan untuk membatasi ekspresi di ruang digital tidak boleh berada di tangan satu entitas tanpa mekanisme kontrol.

Di Amerika Serikat, akuntabilitas dalam moderasi konten dibangun melalui kombinasi antara prinsip hukum federal dan mekanisme internal platform. Pemerintah tidak memiliki lembaga yang secara langsung mengawasi isi konten daring karena *First Amendment* menjamin kebebasan berekspresi dari campur tangan negara. Mekanisme akuntabilitas terutama diatur melalui *Section 230 Communications Decency Act (47 U.S.C. §230)*, yang memberikan kekebalan hukum kepada platform atas konten yang diunggah pengguna, sekaligus mengizinkan mereka untuk melakukan moderasi secara sukarela terhadap materi yang dianggap tidak pantas. Namun, jika sebuah platform menemukan konten yang melanggar hukum pidana, mereka wajib melaporkannya ke *National Center for Missing and Exploited Children (NCMEC)* atau *Department of Justice (DOJ)* dan menyimpan bukti pelaporan hingga satu tahun. Apabila kewajiban ini diabaikan, platform dapat dikenai sanksi administratif atau pidana. Pengguna yang merasa dirugikan oleh tindakan moderasi memiliki hak untuk mengajukan banding melalui mekanisme internal platform atau melalui jalur peradilan. Dengan demikian, sistem akuntabilitas di AS beroperasi berdasarkan transparansi dan tanggung jawab hukum, tanpa intervensi langsung negara terhadap kebijakan moderasi platform.

Di Inggris, *Online Safety Act (OSA) 2023* menempatkan akuntabilitas sebagai prinsip utama dalam tata kelola moderasi konten melalui model *co-regulation* antara pemerintah dan sektor swasta. Platform digital diwajibkan memiliki *duties of care* untuk memastikan sistem moderasi mereka proporsional, transparan, dan menghormati kebebasan berekspresi. Otoritas independen *Ofcom* berperan sebagai pengawas utama yang menerbitkan *Codes of Practice* sebagai panduan teknis bagi platform dalam menilai risiko dan menangani konten ilegal maupun berbahaya bagi anak (United Kingdom, 2024). Ketika ditemukan pelanggaran, *Ofcom* akan terlebih dahulu mengirimkan pemberitahuan resmi dan memberikan waktu bagi platform untuk melakukan perbaikan. Jika platform tidak menindaklanjuti, *Ofcom* dapat mengeluarkan keputusan konfirmasi, menjatuhkan denda hingga 10% dari pendapatan global, atau memerintahkan pemutusan akses terhadap layanan yang bersangkutan. Pengguna yang kontennya dihapus berhak mengajukan keberatan melalui mekanisme internal platform, sementara keputusan regulator dapat diajukan banding ke *Upper Tribunal* (United Kingdom,



2023). Dengan demikian, akuntabilitas dalam sistem Inggris berjalan dua arah: negara memastikan tanggung jawab platform secara hukum, dan publik tetap memiliki ruang untuk menuntut transparansi serta perlindungan atas hak berekspresinya.

Kanada membangun sistem akuntabilitas yang lebih terstruktur melalui rancangan Online Harms Act (Bill C-63), yang menciptakan tiga lembaga baru: *Digital Safety Commission*, *Digital Safety Ombudsperson*, dan *Digital Safety Office*. Ketiganya memiliki fungsi berbeda tetapi saling melengkapi. *Commission* bertugas membuat standar dan melakukan investigasi kepatuhan; *Ombudsperson* menjadi saluran bagi masyarakat untuk mengajukan keluhan atau banding terhadap keputusan platform; sedangkan *Office* mengelola data dan laporan publik untuk menjaga transparansi (Canada, 2024). Model ini mengedepankan partisipasi publik dan mekanisme *independent review*, yang memungkinkan setiap keputusan pemerintah atau platform diuji ulang berdasarkan asas keadilan prosedural.

Sementara itu, Uni Eropa mengatur akuntabilitas melalui Digital Services Act (Regulation (EU) 2022/2065) dengan menekankan prinsip *transparency by design*. Setiap platform wajib menerbitkan laporan transparansi dua kali setahun yang berisi jumlah konten yang dihapus, dasar penghapusannya, serta permintaan dari otoritas negara. Selain itu, pengguna berhak melakukan banding internal dan, jika ditolak, dapat mengajukan keberatan ke *certified out-of-court dispute settlement body* yang disahkan oleh *Digital Services Coordinator* di negara anggota. Platform berskala besar (*Very Large Online Platforms/VLOPs*) bahkan diwajibkan menjalani audit independen setiap tahun oleh lembaga eksternal yang ditunjuk Komisi Eropa, dengan hasil audit yang harus dipublikasikan secara terbuka (European Union, 2022). Melalui kombinasi antara pengawasan administratif, audit independen, dan keterlibatan publik, Uni Eropa menempatkan transparansi sebagai instrumen utama perlindungan kebebasan berekspresi.

Dari perbandingan di atas, dapat disimpulkan bahwa negara-negara demokratis menolak model pengawasan yang bersifat tersentralisasi di bawah kendali pemerintah. Sebaliknya, mereka mengandalkan sistem berlapis yang menggabungkan lembaga independen, kewajiban transparansi, dan hak banding publik untuk menjaga keseimbangan antara keamanan digital dan kebebasan berekspresi.

Secara keseluruhan, pembelajaran dari negara-negara demokratis menunjukkan bahwa moderasi konten yang efektif tidak hanya bergantung pada pengaturan hukum semata, tetapi juga pada desain tata kelola yang menyeimbangkan kebebasan berekspresi, perlindungan kelompok rentan, dan akuntabilitas platform. Pola regulasi yang digunakan oleh berbagai negara tersebut menegaskan bahwa keamanan digital dan kebebasan publik dapat dijaga secara bersamaan



melalui prinsip-prinsip *rights-based*, *proportionality*, dan partisipasi publik, yang menjadi fondasi utama bagi tata kelola digital demokratis.

5. Rekomendasi Untuk Moderasi Konten yang Berperspektif Hak Asasi Manusia

Isu moderasi konten di Indonesia telah bergerak jauh melampaui ranah teknis penghapusan unggahan digital. Ia kini menjadi persoalan politik-hukum yang menyentuh hak paling dasar: kebebasan berekspresi dan hak warga untuk memperoleh informasi. Dari diskusi kelompok terarah (*Focus Group Discussion* atau FGD) mengenai *Kepmenkominfo No. 522 Tahun 2024* yang kami lakukan, tampak bahwa banyak pemangku kepentingan menilai pendekatan pemerintah masih didominasi oleh logika kontrol, bukan perlindungan hak. Negara menempatkan diri sebagai pengendali ruang digital, sementara masyarakat sipil dan pengguna ditempatkan sebagai objek yang harus diawasi, bukan mitra dalam tata kelola informasi.

Hasil FGD yang kami lakukan bersama pemangku kepentingan menyimpulkan perlunya perubahan arah kebijakan menuju kerangka moderasi konten yang lebih sejalan dengan prinsip-prinsip hak asasi manusia. Setidaknya terdapat FGD kami merumuskan tiga rekomendasi utama: (1) revisi terhadap pasal “meresahkan masyarakat”, (2) penerapan model *co-regulation* yang melibatkan masyarakat sipil, dan (3) penerapan transparansi publik melalui laporan berkala tentang penghapusan konten oleh pemerintah dan platform.

5.1 Meninjau Ulang Pasal “Meresahkan Masyarakat”

Salah satu kritik paling kuat dari para peserta FGD adalah terhadap penggunaan istilah “konten yang meresahkan masyarakat” dan “menggangu ketertiban umum” pada pasal 9 ayat 4 huruf b dan pasal 14 ayat 3 huruf c dalam *Permenkominfo No. 5/2020* serta *Kepmenkominfo No. 522/2024*. Dua istilah ini menjadi dasar pemerintah dalam memerintahkan penghapusan konten, tetapi tidak memiliki kriteria hukum yang pasti.

Salah satu manajer perempuan dari lembaga advokasi reformasi hukum menjelaskan bahwa pasal semacam ini bertentangan dengan asas *lex certa* dalam hukum. “Kalimat ‘meresahkan masyarakat’ tidak punya tolok ukur yang bisa diuji. Siapa yang menentukan masyarakat itu resah atau tidak?” ujarnya. Ia menilai, ketiadaan definisi membuat tafsirnya sangat subjektif dan membuka peluang bagi

aparatus administratif untuk menghapus konten yang tidak disukai oleh pihak berkuasa (FGD Kepmen 522, 25/09/2025).

Persoalan ini juga mendapat perhatian dari seorang direktur perempuan dari lembaga yang fokus di bidang hak digital, yang menilai bahwa istilah semacam itu sering digunakan untuk membungkam ekspresi politik atau konten yang mengungkap kekerasan oleh aparat. “Kalimat ‘meresahkan masyarakat’ itu sudah lama dipakai sebagai alat pembungkam. Sekarang hanya bentuknya yang digital,” ujarnya. Ia menekankan perlunya kejelasan parameter agar kebijakan moderasi tidak berubah menjadi bentuk penyensoran yang dilegitimasi oleh hukum (FGD Kepmen 522, 25/09/2025).


Sementara itu, salah satu peneliti di lembaga advokasi reformasi kebijakan publik menegaskan bahwa pemerintah perlu melakukan revisi terhadap pasal-pasal yang membuka peluang tafsir subjektif (FGD Kepmen 522, 25/09/2025). Ia merekomendasikan agar kategori “konten terlarang” dibatasi hanya pada pelanggaran yang jelas secara hukum, seperti pornografi anak, terorisme, atau penipuan digital, dan menghapus kategori yang bersifat moralistik seperti “meresahkan masyarakat”. Apabila kategori tersebut tetap dipertahankan, maka definisinya harus diuji secara publik melalui proses konsultasi multipihak dan ditopang oleh indikator yang dapat diverifikasi.

Pendekatan berbasis prinsip kerugian (*harm principle*) dapat menjadi solusi: pembatasan ekspresi hanya sah jika konten tersebut terbukti secara nyata menimbulkan kerugian bagi individu atau kelompok lain, bukan sekadar menimbulkan ketidaknyamanan atau perbedaan pendapat. Dengan pendekatan ini, moderasi konten akan berpindah dari rezim kontrol menuju rezim perlindungan, selaras dengan komitmen Indonesia terhadap Kovenan Internasional tentang Hak Sipil dan Politik (ICCPR) yang telah diratifikasi sejak 2005.

5.2 Dari Regulasi ke Co-Regulasi: Pelibatan Masyarakat Sipil

Kritik kedua yang mengemuka dalam FGD adalah monopoli negara dalam menentukan standar moderasi konten. Saat ini, seluruh proses penentuan pelanggaran, penghapusan, hingga penjatuhan sanksi berada di bawah kendali Komdigi, tanpa mekanisme pengawasan eksternal. Para peserta FGD berpendapat bahwa tata kelola semacam ini tidak hanya melanggar prinsip *check and balance*, tetapi juga melemahkan kepercayaan publik terhadap legitimasi kebijakan digital.

Seorang pengamat kebebasan digital di Indonesia, menegaskan perlunya model *co-regulation* yang menempatkan masyarakat sipil sebagai mitra seajar



pemerintah. “Moderasi konten itu bukan urusan satu lembaga. Ia harus dinegosiasikan antara pemerintah, platform, dan masyarakat sipil,” ujarnya. Menurutnya, pengalaman dari *Trust Positif* dan SAMAN menunjukkan bahwa tanpa partisipasi publik, regulasi digital cenderung bersifat top-down dan kehilangan kredibilitas sosial (FGD Kepmen 522, 25/09/2025).


Konsep *co-regulation* ini juga didukung oleh salah satu peneliti di lembaga advokasi reformasi kebijakan publik, yang merekomendasikan pembentukan dewan pengawas bersama atau *multistakeholder council* untuk memverifikasi permintaan penghapusan konten (FGD Kepmen 522, 25/09/2025). Dewan ini dapat beranggotakan perwakilan dari pemerintah, platform digital, akademisi, dan organisasi masyarakat sipil yang memiliki kompetensi di bidang kebebasan berekspresi dan teknologi. Dalam praktiknya, dewan semacam ini akan bertugas meninjau laporan konten yang kontroversial, memediasi sengketa antara pemerintah dan platform, serta mengeluarkan rekomendasi publik yang bersifat konsultatif. Ia menilai pelibatan masyarakat sipil bukan hanya persoalan legitimasi demokratis, tetapi juga soal efektivitas kebijakan. “Kalau semua keputusan diambil pemerintah, trust publik akan menurun dan platform jadi bingung harus ikut siapa,” ujarnya. Menurutnya, model *co-regulation* memungkinkan kebijakan digital lebih adaptif terhadap dinamika sosial dan teknologi, karena keputusan dihasilkan melalui dialog antaraktor, bukan sekadar instruksi administratif.

Salah satu manajer perempuan dari lembaga advokasi reformasi hukum menambahkan bahwa pelibatan masyarakat sipil juga memiliki nilai hukum penting: menjadi jaminan partisipasi publik sebagaimana diamanatkan dalam *Undang-Undang Nomor 12 Tahun 2011* tentang Pembentukan Peraturan Perundang-undangan. “Masyarakat sipil harus menjadi bagian dari proses pengawasan, bukan sekadar penonton,” katanya (FGD Kepmen 522, 25/09/2025). Dengan cara itu, pengambilan keputusan tentang penghapusan konten akan lebih transparan dan lebih tahan terhadap intervensi politik.

5.3 Transparansi dan Akuntabilitas Moderasi Konten

Rekomendasi ketiga menyangkut pentingnya transparansi dan keterbukaan data dalam praktik moderasi konten. Selama ini, keputusan pemerintah untuk meminta penghapusan konten dilakukan secara tertutup. Tidak ada laporan publik yang memuat berapa banyak konten yang dihapus, kategori pelanggarannya, atau siapa yang mengajukan permintaan. Ketertutupan ini menciptakan ketegangan antara hak publik untuk tahu dan kepentingan negara untuk mengendalikan.

Salah satu kreator digital di bidang edukasi sejarah menceritakan bagaimana mediana menerima notifikasi dari platform X (Twitter) bahwa salah satu



unggahan mereka dilaporkan oleh Komdigi. “Kami tidak tahu alasan pastinya, tidak ada surat resmi, tidak ada mekanisme keberatan,” ujarnya (FGD Kepmen 522, 25/09/2025). Ia menilai bahwa tanpa transparansi, kebijakan moderasi akan selalu dicurigai sebagai alat politik.

Indonesia perlu mengadopsi model *transparency reporting* seperti yang diterapkan di Uni Eropa melalui *Digital Services Act* (DSA). Dalam model tersebut, baik pemerintah maupun platform diwajibkan mempublikasikan laporan periodik tentang jumlah dan jenis konten yang dihapus, alasan penghapusan, sumber permintaan (negara, individu, atau organisasi), serta hasil banding jika ada. Laporan tersebut harus dapat diakses publik dan diaudit oleh lembaga independen.

Usulan serupa juga muncul dari kalangan jurnalis. Salah satu anggota bidang internet di organisasi jurnalis menyatakan bahwa keterbukaan data *take down* akan membantu publik menilai apakah kebijakan moderasi proporsional. “Kalau datanya terbuka, kita bisa tahu apa yang banyak dihapus—berita, satire, atau ujaran kebencian. Dari situ baru bisa menilai apakah pemerintah bekerja dengan benar,” ujarnya (FGD Kepmen 522, 25/09/2025). Transparansi menjadi instrumen akuntabilitas yang memungkinkan masyarakat mengevaluasi motif dan dampak dari setiap penghapusan konten.

Lebih jauh, salah satu pengamat kebebasan digital di Indonesia mengusulkan agar sistem SAMAN dilengkapi dengan dasbor publik yang menampilkan data *real-time* tentang jumlah permintaan penghapusan konten, kategori pelanggaran, dan lembaga pemohon. Baginya, keterbukaan ini bukan hanya soal akuntabilitas, tetapi juga pendidikan publik. “Kita perlu tahu bagaimana negara bekerja di ruang digital supaya bisa ikut mengawasinya,” ujarnya (FGD Kepmen 522, 25/09/2025). Dengan demikian, transparansi menjadi jembatan yang menghubungkan kepentingan keamanan digital dengan hak publik untuk tahu.

Dengannya, FGD bersama pemangku kepentingan menekankan bahwa arah reformasi kebijakan moderasi konten di Indonesia perlu bergerak dari logika kontrol menuju logika kolaborasi. Revisi terhadap pasal “meresahkan masyarakat” akan memperkuat kepastian hukum dan mencegah penyalahgunaan tafsir subjektif. Penerapan *co-regulation* akan memastikan bahwa masyarakat sipil menjadi bagian dari proses, bukan korban dari keputusan. Dan transparansi melalui laporan publik akan mengembalikan kepercayaan antara pemerintah, platform, dan warga digital.

Dengan tiga langkah tersebut, Indonesia berpeluang membangun sistem moderasi konten yang lebih berpihak pada hak asasi manusia, sebuah tata kelola digital yang tidak hanya menjaga ketertiban, tetapi juga menjamin kebebasan. Karena pada akhirnya, ruang digital yang sehat bukanlah ruang yang steril dari



perbedaan, melainkan ruang yang mampu menampung berbagai suara tanpa rasa takut.



Daftar Pustaka

- Aliansi Jurnalis Independen [AJI]. (2025, July). *Kertas Kebijakan: Kepmen Kominfo 522 Berpotensi Mengancam Kebebasan Pers dan Berekspresi di Indonesia*. Aliansi Jurnalis Independen. <https://aji.or.id/system/files/2025-08/inkertas-kebijakankepmen-kominfo-522final-1.pdf>
- Asosiasi Penyelenggara Jasa Internet Indonesia [APJII]. (2024, February). *APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang*. APJII. <https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang>
- Bambang, J. U. I., Najwa, N., Rahmadani, M. R., Salsabil, H., Sulistyoko, A., Wardani, M., & Muhajir, A. (2025). Kebebasan berbicara di media sosial: Antara regulasi dan ekspresi. *Student Research Journal*, 3(1), 87–96.
- Canada. (2024). *Bill C-63: An Act to enact the Online Harms Act, to amend the Criminal Code, the Canadian Human Rights Act and An Act respecting the mandatory reporting of Internet child pornography by persons who provide an Internet service and to make consequential and related amendments to other Acts*. Ottawa: Parliament of Canada. https://www.parl.ca/Content/Bills/441/Government/C-63/C-63_1/C-63_1.PDF
- Center for Indonesian Policy Studies [CIPS]. (2025a). *Bayang-Bayang Penyensoran: Perkembangan Kebijakan Moderasi Konten Indonesia*. CIPS.
- Center for Indonesian Policy Studies [CIPS]. (2025b). *Evaluasi kebijakan moderasi konten digital di Indonesia*. Jakarta: CIPS.
- European Union. (2022). *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services (Digital Services Act) and amending Directive 2000/31/EC*. Official Journal of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065>
- Galli, F., Loreggia, A., & Sartor, G. (2022). The regulation of content moderation. In *International conference on the legal challenges of the fourth industrial revolution* (pp. 63–87). Cham: Springer International Publishing.
- Humairoh, A., & Nurjannah, A. (2024). Moderasi konten dan literasi digital: Tantangan komunikasi publik di era disinformasi di Indonesia. *Journal Central Publisher*, 2(12), 2891–2898.

Kementerian Komunikasi dan Informatika Republik Indonesia. (2020). *Peraturan Menteri Komunikasi dan Informatika Nomor 5 Tahun 2020 tentang Penyelenggara Sistem Elektronik Lingkup Privat*. Jakarta: Kominfo.

Kementerian Komunikasi dan Informatika Republik Indonesia. (2019). *Peraturan Pemerintah Nomor 71 Tahun 2019 tentang Penyelenggaraan Sistem dan Transaksi Elektronik*. Jakarta: Kominfo.

Kementerian Komunikasi dan Informatika Republik Indonesia. (2024a). *Keputusan Menteri Komunikasi dan Informatika Nomor 172 Tahun 2024 tentang Sistem Kepatuhan Moderasi Konten (SAMAN)*. Jakarta: Kominfo.

Kementerian Komunikasi dan Informatika Republik Indonesia. (2024b). *Keputusan Menteri Komunikasi dan Informatika Nomor 522 Tahun 2024 tentang Perubahan atas Keputusan Menteri Nomor 172 Tahun 2024*. Jakarta: Kominfo.

Kemp, S. (2025, Februari). Digital 2025: Global Overview Report — DataReportal – Global Digital Insights. DataReportal.
<https://datareportal.com/reports/digital-2025-global-overview-report>

SAFENet. (2025a, September 5). Pelanggaran digital selama masa demonstrasi. SAFENet - Southeast Asia Freedom of Expression Network.
<https://safenet.or.id/id/2025/09/pelanggaran-digital-selama-masa-demonstrasi/>

SAFENet. (2025b, September 10). Dokumentasi Kasus TAUD dalam Aksi Agustus-September 2025. SAFENet - Southeast Asia Freedom of Expression Network.
<https://safenet.or.id/id/2025/09/dokumentasi-kasus-taud-dalam-aksi-agustus-september-2025/>


Toptchiyska, D. (2023). Legal Aspects Of Content Moderation On Social Media Platforms: A Comparative Perspective. Available at SSRN 4901501.

Twimbit. (2024). Navigating 2030: AI and the Attention Economy in Indonesia. Twimbit.
<https://cdn.twimbit.com/uploads/2024/03/06073340/Navigating-2030-AI-and-Attention-Economy-Indonesia.pdf>

Undang-Undang Republik Indonesia Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (Lembaran Negara Republik Indonesia Tahun 2008 Nomor 58).

Undang-Undang Republik Indonesia Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (Lembaran Negara Republik Indonesia Tahun 2016 Nomor 251).

United Kingdom. (2023). Online Safety Act 2023. London: The Stationery Office.
<https://www.legislation.gov.uk/ukpga/2023/50/content>



United Kingdom. (2024). Online Safety Act: Illegal Content Codes of Practice 2024 – Explanatory Memorandum. London: GOV.UK.
<https://www.gov.uk/government/publications/online-safety-act-illegal-content-codes-of-practice-2024-explanatory-memorandum>

United States. (n.d.). Amendment I – Freedom of Speech, Press, etc. Washington, DC: U.S. Congress. <https://constitution.congress.gov/constitution/amendment-1/>

United States Code. (n.d.). 18 U.S.C. § 2258A: Reporting requirements of providers. Washington, DC: U.S. Government Publishing Office.
<https://www.law.cornell.edu/uscode/text/18/2258A>

United States Code. (n.d.). 47 U.S.C. § 230: Protection for private blocking and screening of offensive material. Washington, DC: U.S. Government Publishing Office.
<https://www.law.cornell.edu/uscode/text/47/230>

Tentang Remotivi

Didirikan pada tahun 2010, Remotivi berupaya mendorong partisipasi publik yang aktif dan bermakna dalam setiap aspek kehidupan sosial melalui pemenuhan hak komunikasi. Kami percaya perubahan dimulai ketika warga memiliki akses penuh terhadap informasi dan kesempatan untuk didengar.

Kerja-kerja Remotivi dilakukan dalam tiga area:

- Kajian dan advokasi – membuka ruang dialog kritis tentang praktik media dan kebijakan komunikasi
- Literasi warga – membekali publik dengan kemampuan untuk memahami, mengkritisi, dan memanfaatkan media secara bermakna
- Penguatan narasi sipil – mengangkat perspektif dan suara-suara yang sering terpinggirkan dalam ruang publik

Sejak pendiriannya, biaya operasional Remotivi didanai oleh donasi publik, hasil usaha, proyek penelitian, dan hibah lembaga donor.

Tentang Penulis

Muhammad Heychael - Peneliti Utama

Heychael telah bekerja sebagai peneliti media dan komunikasi Selama empat belas tahun. Pada tahun 2011, ia turut mendirikan Remotivi. Kajian Heychael fokus pada keterkaitan antara media, marginalitas, dan demokrasi. Ia secara aktif melakukan penelitian, menulis opini, dan memberikan kuliah serta mengisi seminar mengenai topik-topik seperti komunikasi politik dan gerakan sosial, pemberitaan media tentang terorisme dan konflik keagamaan, kebebasan berekspresi, disinformasi, serta kebijakan dan keamanan media digital.

Yoga Firman Firdaus - Asisten Peneliti

Yoga Firman Firdaus adalah mahasiswa tingkat akhir Jurnalistik di Fikom, Universitas Padjadjaran. Ia aktif menulis kajian dan esai di bidang media serta komunikasi. Ia juga merupakan pendiri Komunitas Peminat Ilmu Komunikasi (Kompilasi), sebuah komunitas yang berfokus pada penulisan, diskusi, dan produksi konten seputar studi komunikasi.

Semua Bisa Bersuara

REMOTIVI

2025

Jl. Raya Kebayoran Lama No.18CD,
RT.004/RW.010, Grogol Selatan,
Kecamatan Kebayoran Lama,
Kota Jakarta Selatan,
Daerah Khusus Ibukota Jakarta 12220

 [remotivi.or.id](https://twitter.com/remotivi.or.id)  [Remotivi.or.id](https://www.youtube.com/Remotivi.or.id)
 [remotivi.or.id](https://www.instagram.com/remotivi.or.id)  www.remotivi.or.id



REMOTIVI
Pusat Kajian Media & Komunikasi