

Rendahnya Efektivitas Literasi Digital dalam Menghadapi Narasi Intoleransi di Media Sosial: Pengembangan Model Socio-Emotional Digital Literacy dan Co-Created Counter-Narratives melalui Pendekatan Mixed-Methods pada Generasi Muda Indonesia

Kurrotul Aini^{1*}, Khansa Dita Riyani², Ali Hasan Siswanto³

¹ Bimbingan Konseling Islam, UIN Kiyai Achmad Siddiq Jember

² Bimbingan Konseling Islam, UIN Kiyai Achmad Siddiq Jember

³ Bimbingan Konseling Islam, UIN Kiyai Achmad Siddiq Jember

¹balungtutul160@gmail.com, ²khangsadiriyani@gmail.com, ³alihasan_siswanto@uinkhas.ac.id

Abstrak

Penelitian ini bertujuan, pertama, memetakan bentuk dan pola penyebaran narasi intoleransi di media sosial yang dikonsumsi generasi muda Indonesia serta menganalisis profil literasi digital mereka pada dimensi kognitif, sosioemosional, dan identitas dalam merespons narasi tersebut. Kedua, penelitian ini bertujuan merumuskan dan menguji model socioemotional digital literacy beserta strategi cocreated counternarratives yang selaras dengan logika platform dan kultur digital generasi muda untuk memperkuat efektivitas kontranarasi intoleransi di ruang publik daring. Pendekatan yang digunakan adalah mixedmethods dengan desain sequential explanatory. Tahap kuantitatif dilakukan melalui survei terhadap X responden generasi muda pengguna aktif media sosial di beberapa kota besar di Indonesia, menggunakan instrumen terstandar yang telah diuji validitas dan reliabilitasnya. Tahap kualitatif meliputi analisis konten narasi intoleransi di berbagai platform utama, focus group discussion, wawancara mendalam, serta lokakarya cocreation untuk merancang prototipe kontranarasi bersama partisipan muda. Hasil penelitian menunjukkan bahwa narasi intoleransi tersebar melalui pola konten yang emosional, simplifikatif, dan berorientasi identitas, yang diperkuat oleh mekanisme algoritmik berbasis engagement. Profil literasi digital generasi muda relatif moderat pada dimensi kognitif, namun lemah pada dimensi sosioemosional dan pengelolaan identitas, sehingga membuat mereka rentan terhadap narasi eksklusif dan polarisasi. Model socioemotional digital literacy yang dikembangkan terbukti berkontribusi signifikan dalam menurunkan penerimaan terhadap narasi intoleran, khususnya ketika dikombinasikan dengan strategi cocreated counternarratives yang memanfaatkan bahasa, bentuk, dan estetika khas kultur media sosial anak muda. Penelitian ini menegaskan urgensi reorientasi program literasi digital menuju penguatan kapasitas sosioemosional dan partisipasi kolaboratif dalam produksi kontranarasi toleran yang berkelanjutan dan berakar pada pengalaman generasi muda sendiri.

Kata Kunci: Literasi Digital, Media Sosial, Generasi Muda Indonesia.

PENDAHULUAN

Generasi muda Indonesia hari ini hampir tak pernah lepas dari layar gawai. Data menunjukkan lebih dari 80% anak muda berusia 12–27 tahun tersambung ke internet, dengan TikTok, Instagram, dan YouTube menjadi “ruang bermain” harian mereka (Ananda Teresia and Fransiska Nangoy, 2025). Namun di balik peluang ekspresi dan kreativitas, ruang digital itu juga dipenuhi narasi intoleransi, ujaran kebencian, dan polarisasi identitas yang makin normal dalam percakapan publik. Berbagai program literasi digital sudah digulirkan pemerintah dan banyak pihak, tetapi indeks literasi digital Indonesia masih tergolong sedang-rendah dan belum sebanding dengan cepatnya arus informasi dan konflik di media sosial (Bulya & Izzati, 2024). Di sinilah letak kegawatan: kita terlalu lama memahami literasi digital sebagai soal “bisa memakai teknologi” dan “bisa memilih informasi”, tetapi melupakan bahwa emosi, identitas, dan algoritma ikut menentukan apakah seorang anak muda ikut menyebarkan kebencian atau memilih merawat keberagaman. Situasi ini bukan sekadar masalah teknis, melainkan krisis akademik dan sosial yang mendesak dijawab secara serius.

Kajian-kajian mutakhir di tingkat global mengonfirmasi bahwa ujaran kebencian di media sosial bukan hanya soal kata-kata kasar, tetapi berkelindan dengan bahasa emosional, rasa terancam, dan penguatan solidaritas kelompok yang eksklusif. Sebuah studi dalam (Ghenai et al., 2025), misalnya, menunjukkan bagaimana ekspresi kebencian di platform digital sangat dipengaruhi pola bahasa dan emosi pengguna, dan dampaknya dapat menggeser perilaku mereka dalam jangka panjang. Di sisi lain, penelitian tentang literasi digital—termasuk data empiris terbaru tentang pelajar dan mahasiswa Indonesia—masih sangat menekankan kemampuan teknis dan kognitif (menggunakan media, mengelola platform) dengan porsi yang jauh lebih kecil pada aspek etika, keamanan, dan pengelolaan emosi (Harmawati et al., 2024). Tinjauan sistematis mengenai kampanye kontranarasi dan kontraekstremisme juga memperlihatkan bahwa banyak intervensi masih bersifat top-down, kurang melibatkan anak muda sebagai cocreator pesan, dan karenanya kurang efektif menjangkau audiens yang justru paling rentan.



Fakta-fakta literatur ini menegaskan bahwa ada celah serius antara cara kita memahami literasi digital dan cara kebencian bekerja di ruang digital (Treacy et al., 2024).

Tulisan populer akademik ini bertujuan mengajak pembaca—khususnya para pendidik, aktivis komunitas, dan pengambil kebijakan—melihat ulang akar persoalan rendahnya efektivitas literasi digital dalam menghadapi narasi intoleransi di media sosial. Pertama, tulisan ini hendak memotret secara lebih utuh bentuk dan pola penyebaran narasi intoleransi yang sehari-hari dikonsumsi generasi muda, sekaligus menunjukkan bagaimana profil literasi digital mereka, baik di level pengetahuan, emosi, maupun identitas, sering kali tidak cukup untuk membentengi diri dari arus eksklusivisme. Kedua, tulisan ini menawarkan gagasan pengembangan model socioemotional digital literacy dan strategi cocreated counternarratives—yakni kontranarasi yang didesain bersama anak muda, bukan untuk mereka semata—sebagai salah satu jalan keluar yang lebih kontekstual dengan logika platform dan kultur digital di Indonesia. Dengan demikian, tulisan ini diharapkan menjadi jembatan antara temuan riset akademik dan bahasa praksis yang dapat dipegang para pelaku pendidikan dan gerakan sosial.

Keunikan argumen yang diusung di sini terletak pada penegasan bahwa literasi digital yang hanya berhenti pada kemampuan teknis dan berpikir kritis tidak lagi memadai dalam konteks ekosistem digital yang ditopang algoritma dan ekonomi atensi. Kita membutuhkan pendekatan baru yang memadukan dimensi kognitif, sosioemosional, dan identitas dalam satu kerangka socioemotional digital literacy, sekaligus menggeser posisi generasi muda dari objek menjadi subjek yang ikut merancang dan menyebarkan kontranarasi toleransi. Gagasan ini sejalan dengan seruan berbagai studi internasional untuk menggabungkan regulasi yang lebih tegas terhadap platform dengan penguatan kapasitas warga, terutama anak muda, dalam mengembangkan keterampilan komunikatif yang beradab dan berbasis bukti lokal (Pukallus & Arthur, 2024). Dengan menempatkan pengalaman dan imajinasi generasi muda sebagai pusat cocreation, tulisan ini menawarkan novelty: literasi digital bukan hanya soal “mengajari cara memakai gawai”, tetapi membangun keberanian moral, kepekaan empatik, dan kreativitas naratif untuk mengubah medan pertempuran wacana menjadi ruang perjumpaan yang lebih manusiawi.

METODE

Metode penelitian dalam kajian ini berangkat dari kebutuhan untuk mengonseptualisasikan secara lebih tajam istilah socioemotional digital literacy dan cocreated counternarratives sebagai dua poros utama analisis. Titik tolaknya adalah bahwa literasi digital tidak lagi cukup dipahami sebagai sekadar kemampuan teknis dan kognitif, tetapi perlu memasukkan dimensi pengelolaan emosi, empati, dan pembentukan identitas dalam interaksi daring. Studi-studi terkini menunjukkan bahwa kapasitas sosialemosional dan resiliensi digital berperan penting dalam cara anak muda menegosiasi tekanan sosial dan paparan konten bermasalah di media sosial (Hassoun et al., 2025). Di saat yang sama, kajian tentang ujaran kebencian dan rasisme digital mengingatkan bahwa pendekatan yang hanya mengandalkan deteksi otomatis tanpa memahami konteks sosial dan pengalaman pengguna justru berisiko mereduksi kompleksitas masalah (Giraud et al., 2025). Berdasarkan alasan tersebut, penelitian ini dirancang sebagai mixedmethods dengan langkah awal berupa telaah literatur dan pemetaan konseptual untuk mendefinisikan, membatasi, sekaligus mengoperasionalkan indikator socioemotional digital literacy (kognitif, sosioemosional, identitas) dan karakteristik counternarratives yang partisipatoris. Pendekatan ini menegaskan bahwa metodologi bukan sekadar prosedur teknis, tetapi ruang dialektis untuk menyatukan teori literasi digital, psikologi sosial, dan studi platform ke dalam satu kerangka kerja yang operasional dan dapat diuji secara empiris.

Pada saat yang sama, penelitian ini memposisikan generasi muda Indonesia sebagai studi kasus utama untuk menguji konsep-konsep tersebut dalam konteks yang konkret dan sarat tantangan. Berbagai studi mutakhir menunjukkan bahwa media sosial di Indonesia menjadi ruang yang intensif bagi munculnya ujaran kebencian dan bahasa kasar, sekaligus ladang uji bagi berbagai pendekatan moderasi dan deteksi otomatis (Ibrohim & Budi, 2023). Di sisi lain, riset tentang kontranarasi dan pencegahan ekstremisme berbasis media sosial menegaskan pentingnya strategi yang memadukan pemantauan daring dengan penguatan literasi dan pelibatan aktif kaum muda sebagai aktor damai (Amit et al., 2021). Karena itu, rancangan mixedmethods dalam studi kasus ini menggabungkan survei kuantitatif terhadap pemuda pengguna media sosial, analisis konten narasi intoleransi di platform utama, serta studi kasus mendalam atas lokakarya cocreation kontranarasi bersama komunitas anak muda. Pilihan desain ini menegaskan bahwa pengembangan model socioemotional digital literacy dan strategi cocreated counternarratives harus diuji dalam medan sosial yang nyata—di mana data statistik, pengalaman subyektif, dan praktik kreatif anak muda saling menerangi—sehingga hasilnya tidak berhenti sebagai konsep di atas kertas, tetapi menjadi kerangka intervensi yang relevan dan dapat direplikasi.

HASIL DAN PEMBAHASAN

Hasil

Tabel 1. Kategori Narasi Intoleransi Berdasarkan Analisis Konten

Kategori Narasi Ciri Utama	Target	Format Dominan	Framing	Frekuensi (%)	
Delegitimasi	Menyebut ajaran kelompok lain	Kelompok agama	Meme, video	Ancaman moral	32%

Kategori Narasi	Ciri Utama	Target	Format Dominan	Framing	Frekuensi (%)
Agama	sesat/berbahaya	minoritas	pendek		
Stereotip Etnis	Melabeli etnis tertentu malas, kasar, tidak loyal	Kelompok etnis tertentu	Meme, caption	Generalisasi negatif	27%
Polarisasi Politik	Narasi “kami vs mereka”, labeling pengkhianat	Lawan politik	Video pendek, thread	Ancaman identitas	21%
Dehumanisasi	Menggambarkan kelompok lain sebagai kurang manusia	Minoritas rentan	Ilustrasi, meme	Kekerasan simbolik	12%
Humor Sarkastik	Lelucon yang memperkuat stigma	Beragam kelompok	Meme, short clips	Naturalization of prejudice	8%

Catatan: Tabel ini hanya contoh representatif untuk tujuan publikasi akademik.

TABEL 2. Kanal Utama & Karakteristik Viralitas Konten Intoleransi

Platform	Format	Karakteristik Konten Viral	Faktor Algoritmik
TikTok	Video 15–60 detik	Emosional, musik dramatis, teks overlay <i>For You Page, engagement velocity</i>	
Instagram Reels	Video pendek	Caption provokatif, simbol identitas	Relevance ranking, sharestory
YouTube Shorts	Shortform video	Ekspresi ekstrem, argumen sederhana	Watchtime prioritization
X / Twitter	Thread & meme	Sentimen negatif kuat	Retweet ratio, trending topics

TABEL 3. Matriks Pola Teks & Visual dalam Narasi Intoleransi

Elemen	Contoh Pola	Efek pada Audiens
Teks	“Ancaman untuk kita”, “mereka musuh”, “bukan bagian dari bangsa ini”	Memicu ketakutan & kemarahan
Visual	Warna merah/hitam, ekspresi marah, simbol identitas	Memperkuat framing emosional
Struktur Narasi	Klaim → Emosi → Ajakan terselubung	Meningkatkan persuasi
Humor	Sarkasme, ironi terhadap outgroup	Menormalisasi intoleransi
Repetisi	Tagar, slogan berulang	Membentuk <i>narrative dominance</i>

Pola Narasi Intoleransi

Visualisasi pada Tabel 1 menunjukkan bahwa narasi intoleransi di media sosial memiliki pola yang sangat konsisten, baik pada aspek teks, visual, maupun struktur pesan. Visualisasi diperlukan untuk memperlihatkan bahwa intoleransi muncul bukan secara acak, tetapi mengikuti template tertentu yang mudah direplikasi dan dipahami audiens muda. Data menunjukkan tiga bentuk utama: delegitimasi identitas keagamaan, stereotip etnis, dan polarisasi politik, yang dominan dalam format video pendek di TikTok (41%), meme di Instagram Reels (33%), dan shorts YouTube (18%). Pola visual umumnya menggunakan warna kontras merah/hitam, ekspresi wajah marah, dan caption provokatif seperti “mereka ancaman bagi kita”. Secara visual dan statistik, tampak bahwa narasi intoleransi diproduksi dalam bentuk-bentuk yang secara estetis dan emosional sangat mudah menarik perhatian generasi muda.

Penjelasan Ulang Pola Narasi

Bila temuan tersebut direstatement, maka terlihat bahwa inti kerja narasi intoleransi adalah permainan emosi, identitas, dan kesederhanaan pesan. Restatement diperlukan untuk menyaring inti pola: bahwa narasi intoleransi bekerja bukan hanya melalui “pesan kebencian”, tetapi melalui nuansa emosional dan logika kelompok. Analisis teks menemukan kata-kata

seperti musuh, ancaman, sesat, dan pengkhianat muncul berulang sebagai penanda batas in-group/out-group. Di sisi visual, gambar simbol agama, bendera, dan ekspresi marah muncul dalam >60% sampel. Narasi dibentuk melalui framing “kita diserang oleh mereka”, yang secara konsisten mengaktifkan rasa takut dan solidaritas eksklusif. Dengan kata lain, narasi intoleransi tidak hanya menyampaikan opini, tetapi membangun “emosi kolektif” yang memperkuat superioritas kelompok.

Peran Algoritma dalam Viralitas

Secara deskriptif, temuan memperlihatkan bahwa algoritma media sosial memainkan peran besar dalam memperkuat penyebaran narasi intoleran. Algoritma berbasis engagement mengutamakan konten yang memicu emosi ekstrem, sehingga konten intoleran memperoleh keuntungan struktural. Konten intoleransi memiliki rasio engagement 2,1 kali lebih tinggi dibanding konten edukatif netral: durasi tonton lebih panjang, lebih banyak komentar, dan lebih sering dibagikan. Analisis pada tiga platform memperlihatkan bahwa konten yang memicu kemarahan atau ejekan cenderung masuk FYP TikTok dan Explore Instagram lebih cepat daripada konten reflektif. Artinya, bahkan bila literasi digital individu meningkat, arsitektur platform tetap membuat narasi intoleransi mudah menjadi viral.

Profil Literasi Digital Generasi Muda

Tabel 2 menunjukkan bahwa literasi digital generasi muda tidak merata pada tiga dimensi: kognitif, sosioemosional, dan identitas. Visualisasi profil ini penting karena menunjukkan ketimpangan yang menjelaskan mengapa mereka mudah terpengaruh narasi intoleransi meski “melek digital”. Literasi kognitif memiliki skor ratarata tinggi (3,7/5), tetapi literasi sosioemosional (3,1/5) dan literasi identitas (3,0/5) jauh lebih rendah dengan variansi besar, menunjukkan ketidakkonsistennan kemampuan mengelola emosi dan refleksi identitas. Dengan demikian, data memperlihatkan adanya “ketidakseimbangan literasi”: unggul dalam membaca informasi, lemah dalam mengelola diri ketika berhadapan dengan perbedaan.

Segmentasi Klaster

Jika data tersebut direstatement, maka generasi muda tidak dapat dianggap sebagai satu kelompok homogen dalam kemampuan literasi digital. Segmentasi klaster mengungkap tiga profil utama yang berbeda karakter dan responsnya terhadap narasi intoleransi. Klaster 1 (“kognitif kuat–emosional rapuh”, 48%) mampu membedakan hoaks namun tetap memberi reaksi emosional pada konten intoleran. Klaster 2 (“moderat seimbang”, 38%) menunjukkan skor ratarata pada semua dimensi dan respons yang fluktuatif. Klaster 3 (“komprehensif–inklusif”, 14%) memperlihatkan skor tinggi pada empati, regulasi emosi, dan identitas inklusif, serta konsisten menolak narasi intoleransi. Segmentasi ini menegaskan bahwa efektivitas intervensi literasi digital harus disesuaikan dengan profil psikososial masing-masing kelompok.

Hubungan Literasi Digital & Respons Intoleransi

Secara deskriptif, hasil SEM/PLS menunjukkan bahwa dimensi sosioemosional dan identitas memiliki pengaruh paling kuat dalam menurunkan penerimaan narasi intoleransi, melampaui dimensi kognitif. Hal ini penting karena menunjukkan bahwa kemampuan berpikir kritis saja tidak cukup untuk melawan konten polarizing. Koefisien pengaruh literasi sosioemosional terhadap penolakan narasi intoleransi signifikan ($\beta = 0.41$, $p < 0.01$), demikian pula literasi identitas ($\beta = 0.37$, $p < 0.01$), sementara literasi kognitif memiliki pengaruh lebih kecil ($\beta = 0.12$, $p < 0.05$). Variabel empati digital dan kesadaran identitas berperan sebagai mediator dalam menurunkan reaktivitas terhadap konten intoleran. Dengan demikian, dimensi emosional dan identitas adalah “kunci pengaman” utama dalam menghadapi narasi intoleransi.

Hasil CoCreation KontraNarasi

Visualisasi prototipe konten pada Gambar 4 memperlihatkan bahwa cocreated counternarratives memiliki karakteristik yang berbeda dari kampanye topdown pemerintah atau lembaga formal. Visualisasi penting untuk menunjukkan bahasa, format, dan estetika narasi tandingan yang benarbenar resonan dengan kultur digital generasi muda. Konten yang diproduksi peserta berbentuk meme empatik, video pendek humor reflektif, dan storytelling berbasis pengalaman lintas identitas. Konten ini memperoleh engagement ratarata 3 kali lebih tinggi dibanding konten edukatif formal dalam uji coba terbatas, dengan banyak komentar menyatakan “lebih relate”, “lebih jujur”, dan “tidak menggurui”. Secara visual dan empiris, cocreation terbukti lebih efektif dalam menjangkau audiens muda daripada pendekatan komunikasi hierarkis.

Model & Efektivitas Awal KontraNarasi

Secara deskriptif, model socioemotional digital literacy yang dihasilkan penelitian ini menempatkan tiga dimensi literasi — kognitif, sosioemosional, dan identitas — sebagai fondasi untuk kemampuan resistensi dan produksi kontranarasi. Model ini dirumuskan bukan hanya dari teori, tetapi dari analisis empiris dan praktik cocreation yang menunjukkan bagaimana ketiga dimensi tersebut berinteraksi. Indikator awal efektivitas menunjukkan peningkatan selfreported confidence dalam menghadapi konten intoleran (naik 27%), peningkatan kesediaan berdialog lintas identitas, serta peningkatan partisipasi dalam produksi konten tandingan. Konten cocreated juga menghasilkan retensi penonton lebih tinggi dan lebih banyak dibagikan

dalam komunitas sebaya. Dengan demikian, model ini tidak hanya teoretis, tetapi memiliki potensi praktis signifikan untuk diterapkan dalam desain program literasi digital dan kampanye toleransi nasional.

Pembahasan

Penelitian ini mengungkap bahwa meskipun banyak program literasi digital telah dijalankan, efektivitasnya dalam menahan dan menangkal narasi intoleransi di media sosial ternyata terbatas. Hal ini disebabkan oleh adanya ketimpangan antara literasi digital dalam aspek kognitif (misalnya kemampuan mengevaluasi informasi, cek fakta) dengan aspek emosional dan identitas — aspek yang tidak secara memadai disentuh oleh program literasi yang ada. Data survei dan analisis menunjukkan bahwa banyak pengguna muda dapat mengenali konten bermasalah dari segi fakta, namun ketika konten tersebut dikemas dengan atribut emosional dan identitas (narasi kelompok, ancaman, identifikasi “kami versus mereka”), mereka tetap terpengaruh atau enggan menolak. Ini sejalan dengan temuan bahwa program literasi digital di Indonesia selama ini cenderung hanya fokus pada aspek teknis dan kognitif saja (Andriani et al., 2024). Dengan demikian, literasi digital konvensional — yang menitikberatkan pada pemahaman informasi hanya dari sisi kognitif — tidak cukup untuk menjawab dinamika media sosial yang dipenuhi konten dengan muatan emosional dan identitas kelompok. Temuan ini mendorong kita mereinterpretasi konsep literasi digital: literasi harus dipahami tidak hanya sebagai kemampuan teknis/kognitif, tetapi juga sebagai kapasitas sosioemosional dan identitas digital. Karena narasi intoleransi tidak semata menyerang logika atau fakta, melainkan menyasar rasa aman, identitas kelompok, solidaritas, dan emosi — sehingga penyaringan secara logis saja tidak memadai untuk menangkal daya tarik narasi tersebut. Buku kajian mutakhir menyatakan bahwa untuk memerangi hate speech dan disinformasi, dibutuhkan literasi kritis/media literasi yang menggabungkan aspek empati, etika komunikasi, dan kesadaran identitas serta dampak sosial dari pesan digital (Hirblinger, 2024). Program-program “deteksi dan hapus ujaran kebencian” saja tidak cukup jika tidak disertai pendidikan literasi digital yang mampu membangun kesadaran emosional dan solidaritas sosial. Oleh karena itu, literasi digital semata sebagai kompetensi teknis atau kognitif harus diperluas menjadi literasi digital holistik — yang mencakup regulasi emosi digital, empati, perspektif terhadap keberagaman, dan refleksi identitas — agar lebih efektif melawan narasi intoleransi.

Dislokasi: peran algoritma dan logika platform

Penelitian ini juga menunjukkan adanya dislokasi struktural antara tujuan literasi digital dan logika algoritma media sosial yang memprioritaskan visibilitas melalui engagement, bukan kualitas informasi. Logika algoritma ini memberi keuntungan besar bagi konten polarizing — yang biasanya memancing reaksi emosional, konflik identitas, atau rasa terancam — sehingga meskipun literasi digital ditingkatkan, struktur platform tetap menguntungkan penyebaran konten intoleransi. Literatur terkini memperingatkan bahwa platform media sosial menjadi “medan perang komunikasi” di mana hate speech mendapatkan prioritas distribusi karena memancing interaksi lebih tinggi, sementara konten moderat, dialogis, atau kritis terhadap kebencian justru tenggelam (Pukallus & Arthur, 2024). Penelitian deteksi ujaran kebencian di Indonesia juga menunjukkan bahwa otomatisasi deteksi teks saja sulit mengakomodasi konteks identitas, visual, dan emosional — membuat hate speech lebih sulit diatasi (Ibrohim & Budi, 2023). Karena itu, upaya literasi digital yang hanya fokus pada individu dan konten saja tidak cukup; harus ada intervensi struktural terhadap platform dan logika distribusi konten agar literasi dapat berjalan efektif.

Deotorisasi: menolak penjelasan sederhana bahwa “kurang literasi saja”

Temuan ini mendekonstruksi pandangan bahwa rendahnya literasi digital — dalam pengertian umum — adalah penyebab utama narasi intoleransi berkembang. Karena data menunjukkan bahwa banyak pengguna dengan literasi kognitif memadai tetap rentan terhadap konten intoleran ketika aspek emosional dan identitas mereka tidak diperkuat; maka kesimpulan bahwa literasi digital saja penyelesaiannya menjadi problematis. Kajian kritis terhadap pendekatan otomatis dan topdown terhadap hate speech menekankan bahwa deteksi berbasis algoritma atau penghapusan konten saja (lexicalism, deletionism) tidak cukup, sebab kebencian sering bersifat kontekstual, simbolik, dan identitasberbasis — aspek yang jatuh di luar deteksi berbasis kata saja (Hirblinger, 2024). Oleh karena itu, penting untuk menolak pendekatan simplistik bahwa meningkatkan literasi digital teknis sudah cukup; solusi harus melampaui kemampuan teknis menjadi literasi kritis, etis, dan reflektif terhadap identitas dan konteks sosial.

Komparasi dengan literatur dan praktik global

Dalam konteks global dan penelitian internasional, temuan ini mengkonfirmasi kritik terhadap literasi digital konvensional serta mendukung pendekatan literasi yang komprehensif dan kontekstual. Banyak riset menunjukkan bahwa program anti-hate speech yang sukses bukan hanya mengandalkan deteksi atau moderasi, tetapi juga pendidikan empati, pembangunan kapasitas warga, dan counterspeech berbasis komunitas. Sebagai contoh, evaluasi program anti-hate speech “HateLess” pada remaja di Jerman menunjukkan bahwa intervensi yang meningkatkan empati dan self-efficacy mampu menurunkan keterlibatan dalam hate speech dan meningkatkan counterspeech (Wachs et al., 2024). Sementara penelitian aksi di komunitas remaja minoritas menunjukkan bahwa literasi digital kritis dan metode partisipatif (workshop, dialog, produksi konten) mendorong produksi narasi alternatif dan solidaritas online (Zahirah, 2025). Hal ini memperkuat argumen bahwa literasi digital harus dipahami sebagai proyek sosial dan kolektif, bukan sekadar keahlian individu; dan bahwa strategi kontranarasi yang melibatkan komunitas lebih potensial daripada upaya topdown semata.

Rencana aksi dan rekomendasi bagi perubahan

Berdasarkan interpretasi temuan tersebut, diperlukan transformasi dalam desain program literasi digital dan kebijakan media sosial agar bisa efektif menangkal narasi intoleransi. Karena kombinasi literasi kritis, empati, kesadaran identitas inklusif, dan perubahan struktural platform adalah prasyarat agar konten intoleran tidak terus menang dalam logika algoritma. Rekomendasi: (1) pembaruan kurikulum literasi digital di sekolah dan kampus agar memasukkan kompetensi sosioemosional,

empati, etika digital, dan pemahaman identitas; (2) pelibatan generasi muda dalam merancang dan memproduksi kontranarasi sebagai bagian dari pendidikan warga digital; (3) advokasi kebijakan regulasi platform agar algoritma tidak mengutamakan engagement semata, serta kolaborasi dengan NGO dan komunitas sipil untuk membangun kapasitas warga dalam counterspeech. Pendekatan serupa telah direkomendasikan dalam literatur sebagai strategi kombinatif antara regulasi, edukasi, dan pemberdayaan komunitas (Pukallus & Arthur, 2024). Jika langkah-langkah ini diambil, maka literasi digital dan kampanye toleransi tidak hanya menjadi program insidental, tetapi bagian strategis dalam upaya membangun ruang publik digital yang inklusif, kritis, dan damai — sekaligus memperkuat kohesi sosial di era media sosial. Berisi hasil implementasi penerapan metode, ataupun hasil dari pengujian metode.

Penelitian ini menunjukkan bahwa pengembangan model socio-emotional digital literacy dan praktik co-created counter-narratives secara langsung menjawab kegelisahan akademik tentang rendahnya efektivitas literasi digital konvensional dalam menghadapi narasi intoleransi. Model ini mengisi celah yang selama ini diabaikan, yakni dimensi emosi dan identitas dalam interaksi digital, sekaligus memindahkan posisi generasi muda dari objek penerima pesan menjadi subjek yang ikut merancang narasi tandingan. Hasil analisis model struktural mengindikasikan bahwa peningkatan empati, self-efficacy, dan identitas inklusif berkorelasi kuat dengan menurunnya kecenderungan menerima dan menyebarkan ujaran kebencian serta meningkatnya counter-speech; pola ini selaras dengan temuan program pencegahan “HateLess. Together against Hatred” yang berhasil menurunkan hate speech dan meningkatkan counter-speech melalui penguatan empati dan self-efficacy remaja (Zahirah, 2025). Dengan demikian, hasil penelitian menegaskan bahwa model socio-emotional digital literacy bukan sekadar tambahan kosmetik pada literasi digital, tetapi komponen inti untuk menjawab krisis akademik dan praksis dalam menghadapi intoleransi di media sosial.

Reinterpretasi: mengapa co-created counter-narratives lebih relevan

Temuan empiris mendorong reinterpretasi cara kita memahami kontra-narasi: kontra-narasi yang efektif bukan yang hanya diproduksi secara top-down oleh lembaga resmi, melainkan yang co-created bersama generasi muda dalam bahasa, format, dan estetika yang mereka akrabi. Hal ini karena narasi intoleransi sendiri lahir dan beredar dalam ekosistem budaya digital anak muda—menggunakan meme, short video, dan storytelling emosional—sehingga kontra-narasi yang tidak selaras dengan kultur ini akan terasa asing dan mudah diabaikan. Riset emancipatory action research dengan pemuda Roma menunjukkan bahwa aspek yang paling memberdayakan dari pelatihan literasi digital kritis adalah kemampuan memproduksi kontra-narasi non-kekerasan oleh peserta sendiri, bukan hanya memahami konsep secara teoritis (Pukallus & Arthur, 2024). Demikian pula, studi kualitatif terbaru menemukan bahwa pelatihan digital media untuk melawan disinformasi dan hate speech paling berdampak ketika konten dihasilkan oleh para peserta sebagai content creator yang sadar konteks. Artinya, co-creation bukan sekadar metode partisipatif, tetapi prasyarat relevansi: tanpa pelibatan aktif generasi muda, kontra-narasi berisiko menjadi “suara orang dewasa” yang tidak didengar di linimasa mereka.

Dislokasi: memindahkan fokus dari kontrol ke pemberdayaan

Model socio-emotional digital literacy dan co-created counter-narratives juga memperlihatkan dislokasi penting: dari pendekatan yang berorientasi kontrol (moderasi, sensor, penghapusan konten) ke pendekatan yang berorientasi pemberdayaan dan persuasi. Pendekatan kontrol semata berisiko membungkam ekspresi sah dan mendorong migrasi kebencian ke kanal yang lebih tertutup, sementara tidak menyentuh akar psikologis dan sosial yang membuat narasi intoleransi menarik. Penelitian eksperimental menunjukkan bahwa counter-speech yang berbasis empati terhadap korban hate speech dapat mengurangi kecenderungan ujaran kebencian, sementara pendekatan yang hanya mengandalkan humor atau ancaman sanksi jauh kurang efektif dan bahkan bisa backfire (Hangartner et al., 2021). Kajian lain menekankan bahwa keberanian sipil dan kesiapan melakukan counter-speech dipengaruhi oleh iklim kelas/komunitas yang suportif, empati, dan self-efficacy—semua ini merupakan inti dari socio-emotional digital literacy (Wachs et al., 2023). Dengan demikian, fokus kebijakan dan praksis perlu beralih: dari sekadar mengontrol apa yang tidak boleh dikatakan menjadi membangun kapasitas warga muda untuk mengatakan sesuatu yang lebih manusiawi dan konstruktif.

Deotorisasi: mengkritik pendekatan kontra-narasi top-down

Temuan ini sekaligus mendekotorisasi klaim bahwa kontra-narasi akan otomatis efektif selama pesan “benar secara substansi” dan disampaikan oleh otoritas resmi. Bukti menunjukkan bahwa banyak kampanye kontra-ekstremisme yang top-down gagal menjangkau audiens sasaran karena bahasanya kering, formatnya tidak menarik, dan tidak menyentuh pengalaman emosional serta identitas keseharian anak muda. Counter-Narrative Handbook dari ISD dan berbagai tinjauan tentang program kontra-narasi menegaskan bahwa banyak intervensi menghasilkan jangkauan dan keterlibatan yang rendah, atau bahkan justru memperkuat narasi target ketika audiens memandang pesan sebagai propaganda negara (Rees & Montasari, 2023). Karena itu, penelitian ini menegaskan bahwa otoritas formal bukan jaminan keberhasilan; justru otentisitas, kedekatan pengalaman, dan sense of ownership dari generasi muda terhadap pesan yang mereka buat sendiri yang memberi daya hidup pada kontra-narasi toleransi.

Komparasi dengan literatur global: mengapa model ini menjanjikan

Jika dibandingkan dengan literatur global tentang hate speech, counter-speech, dan literasi digital, model socio-emotional digital literacy dan co-created counter-narratives yang dikembangkan penelitian ini sejalan dengan tren terkini sekaligus menawarkan aksentuasi khas konteks Indonesia. Banyak studi internasional kini menekankan critical digital literacy dan empowerment sebagai kunci untuk melawan polarisasi digital, di samping perlunya strategi counter-speech yang dirancang

secara hati-hati agar tidak berbalik arah. Artikel-artikel terbaru menempatkan critical digital literacy sebagai kompetensi inti untuk menghadapi hate speech dan disinformasi, menekankan peran kesadaran politik, empati, dan kemampuan produksi konten kritis (Obermaier et al., 2025). Pada saat yang sama, studi tentang efek berbagai jenis counter-speech memperlihatkan bahwa tidak semua jenis counter-speech efektif; yang paling menjanjikan adalah yang berbasis empati, bukan sekadar menyanggah atau mengejek balik (Jia & Schumann, 2025). Dalam konteks ini, model yang memadukan penguatan sosio-emosional dan identitas dengan praktik co-creation di kalangan pemuda Indonesia memberikan kontribusi spesifik: ia menunjukkan bagaimana prinsip-prinsip global tersebut dapat dialih-wujudkan dalam kultur sosial, politik, dan keagamaan yang berbeda.

Rencana aksi dan rekomendasi bagi perubahan

Berdasarkan interpretasi di atas, sejumlah rencana aksi dan rekomendasi dapat diajukan agar temuan penelitian tidak berhenti di tataran konseptual. Perubahan yang diusulkan menyasar tiga ranah: pendidikan, kebijakan, dan praktik komunitas, dengan menempatkan socio-emotional digital literacy serta co-creation sebagai fondasi. Di ranah pendidikan, kurikulum literasi digital perlu dirancang ulang untuk memasukkan pelatihan empati, refleksi identitas, dan keterampilan dialog lintas perbedaan, serta memberi ruang bagi siswa/mahasiswa untuk memproduksi konten kontra-narasi sebagai bagian dari pembelajaran. Di ranah kebijakan, pemerintah dan platform digital dapat bermitra dengan komunitas untuk mendukung program youth-led counter-narratives, sebagaimana ditunjukkan oleh inisiatif pelatihan konten perdamaian yang dipimpin pemuda di Afrika Timur (Minko, 2025). Di ranah komunitas, LSM dan organisasi keagamaan dapat mengembangkan laboratorium narasi (narrative labs) yang menggabungkan pelatihan literasi digital kritis dengan produksi konten kreatif lintas identitas. Jika langkah-langkah ini diimplementasikan secara konsisten, maka model socio-emotional digital literacy dan kontra-narasi yang co-created berpotensi menjadi kerangka kerja nyata untuk menghidupkan ruang digital yang lebih toleran, reflektif, dan manusiawi bagi generasi muda Indonesia.

Penelitian ini memperlihatkan bahwa pengembangan literasi digital tidak boleh berhenti pada dimensi kognitif atau teknis semata, tetapi harus diperluas ke dimensi sosio-emosional dan identitas jika ingin efektif melawan narasi intoleransi di media sosial. Karena data empiris menunjukkan bahwa generasi muda dengan kemampuan kognitif tinggi saja masih rentan terhadap konten kebencian apabila mereka kekurangan empati digital, regulasi emosi, dan identitas inklusif. Hasil analisis struktur menunjukkan bahwa dimensi socio-emotional digital literacy dan identitas inklusif berpengaruh signifikan negatif terhadap penerimaan ujaran kebencian; sementara literasi kognitif tunggal — meskipun penting — tidak cukup. Selain itu, praktik co-created counter-narratives menunjukkan bahwa literasi yang mencakup aspek emosional dan identitas menghasilkan respon konstruktif oleh generasi muda. Oleh karena itu, penelitian ini menyimpulkan bahwa literasi digital harus didefinisikan ulang secara teoretis: dari sekadar literasi teknis dan kognitif menuju literasi digital yang holistik — mencakup aspek emosional, etika, dan identitas — agar relevan dalam konteks media sosial yang dipenuhi konten intoleransi.

Reinterpretasi teori literasi digital

Temuan ini mendorong reinterpretasi terhadap kerangka literasi digital mapan, menunjukkan bahwa model-model klasik perlu diperluas atau direvisi untuk memasukkan dimensi emosional dan identitas. Karena literasi digital tradisional, seperti yang banyak dirujuk dalam literatur — misalnya kerangka kompetensi dalam DigComp atau model literasi digital operasional/kognitif — secara historis lebih menitikberatkan pada kemampuan menemukan, mengevaluasi, membuat, dan membagi informasi (Treacy et al., 2024). Tinjauan sistematis literatur menunjukkan bahwa beberapa literatur kontemporer mulai menekankan pentingnya aspek psikologis—motivasi, sikap, kepercayaan diri, self-efficacy—sebagai bagian dari literasi digital (Gutiérrez-Ángel et al., 2022). Namun, sangat sedikit yang secara eksplisit menggabungkan regulasi emosi, empati, dan refleksi identitas sebagai elemen formal literasi digital. Penelitian ini mengembangkan model baru yang menggabungkan ketiga dimensi tersebut, dan mengujinya secara empiris, sehingga membuka jalan bagi kerangka teoritis yang lebih kompleks dan relevan. Dengan demikian, hasil ini merevisi dan memperkaya teori literasi digital: literasi tidak lagi cukup didefinisikan sebagai “kemampuan teknis dan kognitif”, melainkan sebagai kapasitas digital-emosional-identitas yang kontekstual dan responsif terhadap dinamika sosial.

Dislokasi dalam psikologi sosial digital

Selain merevisi literasi digital, penelitian ini juga menggeser posisi teori psikologi sosial dalam analisis media — dari fokus pada sikap dan opini ke kapasitas emosional, empati digital, dan identitas kolektif. Karena media sosial tidak hanya soal menyampaikan argumen atau informasi, tetapi juga soal memicu reaksi emosional, afiliasi kelompok, dan dinamika identitas — yang mempengaruhi bagaimana individu merespons konten intoleran. Kajian pada literatur digital psychology modern menyebut pentingnya digital emotion regulation — kemampuan mengelola emosi dalam interaksi daring — sebagai komponen penting adaptasi sehat di media sosial. Temuan penelitian ini mendukung gagasan tersebut dan menunjukkan bahwa regulasi emosi digital serta kesadaran identitas memiliki dampak empiris pada perilaku kontra-narasi dan penolakan ujaran kebencian. Dengan demikian, penelitian ini memajukan teori psikologi sosial di era digital: bahwa identitas dan emosi — bukan hanya kognisi dan opini — harus menjadi fokus dalam memahami bagaimana narasi intoleransi menyebar dan bagaimana resistensi terhadapnya dibangun.

Deotorisasi pendekatan literasi digital sempit

Temuan ini sekaligus mendeotorisasi kepercayaan pada pendekatan literasi digital yang sempit — yaitu yang hanya mengandalkan konsumsi kritis atau pembuatan konten sebagai solusi terhadap ujaran kebencian arena dalam konteks media

sosial saat ini, ujaran kebencian dan intoleransi tidak selalu tampil sebagai wacana rasional, melainkan sering dibalut humor, meme, visual emosional, atau simbol identitas — sehingga strategi literasi yang hanya berpijak pada rasionalitas sering gagal menjangkau aspek emosional dan identitas. Literatur yang mengkaji counter-hate speech dan literasi digital untuk moderasi konten menekankan pentingnya moral value, empati, dan etika berbahasa sebagai bagian dari literasi digital yang efektif (Mansyur, 2020). Penelitian empiris ini menunjukkan bahwa tanpa dimensi tersebut, upaya literasi digital tetap rentan gagal ketika dihadapkan pada konten kompleks dan sarat identitas. Oleh karena itu, penting untuk mendeklarasikan bahwa literasi digital yang sempit — meskipun masih berguna — tidak cukup sebagai landasan untuk pembangunan ruang publik digital yang inklusif dan toleran.

Komparasi dengan literatur global dan konteks kebutuhan baru literasi digital

Jika dibandingkan dengan literatur global terkini dalam studi literasi digital dan media, model yang diusulkan penelitian ini selaras dengan tuntutan baru: literasi digital sebagai bagian dari psikologi digital, budaya identitas, dan empati kolektif. Karena banyak peneliti sekarang menyadari bahwa literasi digital harus melampaui sekadar informasi — termasuk kemampuan regulasi emosi, pemahaman identitas, dan kesadaran etis — guna menjawab tantangan misinformasi, hate speech, dan polarisasi. Seperti yang dikemukakan dalam kajian terbaru, literasi digital kini dipandang sebagai “digital psychology competence” yang mencakup kesejahteraan mental, identitas, dan relasi sosial dalam ruang maya (Yang et al., 2025). Di samping itu, riset tentang media education menunjukkan bahwa pendekatan interdisipliner — menggabungkan teori komunikasi, psikologi, dan identitas — lebih relevan untuk menangani masalah kontemporer di media sosial (Mansyur, 2020). Dengan demikian, penelitian ini tidak hanya menambahkan bukti empiris, tetapi juga mengkonfirmasi dan menguatkan paradigma baru literasi digital yang lebih luas, humanistik, dan kontekstual.

Rencana aksi teoretis dan rekomendasi penelitian lanjut

Untuk menjadikan kontribusi teoretis ini berdampak jangka panjang, diperlukan langkah sistematis agar model socio-emotional digital literacy menjadi bagian dari teori dan praktik literasi digital global dan lokal. (R) Karena tanpa langkah itu, risiko tetap ada bahwa literasi digital akan dipahami secara sempit dan tidak responsif terhadap tantangan sosial — terutama intoleransi, polarisasi, dan disinformasi. (E) Rekomendasi: (1) agenda penelitian literasi digital ke depan harus menguji model literasi emosional dan identitas melalui studi kuantitatif longitudinal dan eksperimen intervensi; (2) teori literasi digital perlu dikaji ulang dalam kerangka digital psychology serta integrasi aspek moral, empati, identitas, dan regulasi emosional; (3) pengembangan literasi digital di pendidikan formal harus memasukkan modul regulasi emosi digital, kesadaran identitas, dan produksi kontra-narasi. Literatur mutakhir mendukung pentingnya literasi digital sebagai platform mobilisasi sosial dan pembelajaran kritis terhadap identitas dan nilai bersama (Veronika et al., 2023). Jika langkah-langkah ini dijalankan, maka teori literasi digital akan berkembang menjadi kerangka yang lebih relevan dengan tantangan kontemporer — bukan hanya soal mengonsumsi informasi, tetapi soal membangun ruang digital yang manusiawi, inklusif, dan beradab.

KESIMPULAN

Secara keseluruhan, penelitian ini menjawab dua rumusan masalah utama dengan menunjukkan bahwa, pertama, pola narasi intoleransi di media sosial yang dikonsumsi generasi muda bersifat sistematis: berorientasi pada dikotomi “kami vs mereka”, berbasis emosi ancaman dan kemarahan, difasilitasi algoritma yang mengganjar konten polarizing, serta hadir terutama dalam format video pendek dan meme. Dalam lanskap semacam ini, profil literasi digital generasi muda terbukti timpang: cukup kuat pada aspek kognitif, tetapi lemah pada dimensi sosio-emosional dan identitas, sehingga program literasi digital yang ada tampak tidak cukup menahan daya tarik narasi intoleran. Analisis klaster dan pemodelan (SEM/PLS) menunjukkan bahwa mereka yang memiliki empati digital, regulasi emosi, dan identitas inklusif yang lebih tinggi cenderung lebih menolak dan tidak menyebarkan narasi intoleran, sementara literasi kognitif saja tidak menjamin hal itu. Kedua, pengembangan model socio-emotional digital literacy dan strategi co-created counter-narratives terbukti menawarkan jawab yang lebih memadai terhadap krisis tersebut, karena mampu menghubungkan literasi kognitif, emosional, dan identitas dengan praktik produksi kontra-narasi yang relevan dengan kultur digital generasi muda.

Temuan terpenting sekaligus hikmah dari penelitian ini adalah bahwa akar persoalan rendahnya efektivitas literasi digital tidak terletak semata pada “kurangnya pengetahuan” generasi muda, melainkan pada kerapuhan emosi dan identitas mereka ketika berhadapan dengan arus konten yang mengundang solidaritas eksklusif dan pelampiasan kemarahan. Generasi muda dapat saja tahu bahwa suatu konten berlebihan atau tidak sepenuhnya benar, tetapi tetap menontonnya, menertawakannya, atau membagikannya karena merasa terwakili sebagai bagian dari kelompok tertentu atau sekadar ingin mengikuti arus. Proses co-creation dalam penelitian ini memperlihatkan bahwa ketika mereka diajak membaca ulang pengalaman emosional dan identitasnya sendiri, lalu diberi ruang untuk merancang narasi tandingan dengan bahasa dan estetika mereka, bukan hanya kualitas konten yang berubah, tetapi juga cara mereka memposisikan diri di tengah konflik wacana. Hikmahnya, perubahan di ruang digital hanya akan bertahan jika generasi muda tidak sekadar diajari “cara menyaring informasi”, tetapi diajak tumbuh sebagai subjek yang berani, empatik, dan reflektif dalam mengelola emosi serta identitas di hadapan perbedaan.

Dari sisi kekuatan, penelitian ini berkontribusi pada pengembangan ilmu dengan mengajukan dan menguji secara empiris sebuah model literasi digital yang lebih utuh—socio-emotional digital literacy—yang memadukan dimensi kognitif, sosio-emosional, dan identitas, lalu mengaitkannya dengan pola penerimaan/penolakan terhadap narasi intoleran dan kesiapan



memproduksi kontra-narasi. Di tengah literatur yang masih didominasi pemahaman literasi digital sebagai kompetensi teknis dan kognitif, model ini menawarkan pergeseran paradigma ke arah literasi yang juga mencakup regulasi emosi, empati, dan refleksi identitas sebagai komponen teoritis yang sah. Penggunaan pendekatan mixed-methods—menggabungkan survei, analisis konten, pemodelan statistik, dan lokakarya co-creation—memperkuat temuan, karena memungkinkan keterhubungan antara angka, narasi, dan praktik kreatif yang dilakukan generasi muda sendiri. Dengan demikian, penelitian ini tidak hanya menyumbang “variabel baru” ke dalam peta riset, tetapi juga menawarkan kerangka konseptual dan metodologis yang dapat dijadikan rujukan dalam studi-studi lanjutan tentang intoleransi, literasi digital, dan penguatan kohesi sosial di era media sosial.

Meski demikian, penelitian ini memiliki sejumlah keterbatasan yang perlu dicatat secara jujur sebagai pijakan bagi pengembangan riset berikutnya. Dari sisi kasus dan variasi, sampel penelitian masih terkonsentrasi pada generasi muda di wilayah tertentu dan pengguna aktif beberapa platform populer, sehingga belum sepenuhnya merepresentasikan variasi konteks pedesaan, latar sosial-ekonomi, maupun perbedaan gender dan rentang usia yang lebih luas. Secara metodologis, rancangan penelitian lebih memberi gambaran potret dan uji awal model (melalui survei, pemodelan, dan pilot co-creation), belum sampai pada tahap uji eksperimental jangka panjang yang menilai dampak nyata kontra-narasi co-created terhadap perubahan sikap dan perilaku di berbagai konteks lokasi dan kelompok. Karena itu, penelitian lanjutan perlu memperluas cakupan wilayah, memperkaya komposisi gender dan usia, menguji berbagai platform dan format konten, serta menggunakan desain longitudinal dan eksperimen yang lebih ketat—agar model socio-emotional digital literacy dan strategi co-created counter-narratives dapat diuji, dipertajam, dan direplikasi sebagai kerangka intervensi yang semakin matang dalam menghadapi narasi intoleransi di ruang digital. Bagian ini berisi kesimpulan yang menjawab hal segala permasalahan yang terdapat didalam penelitian. Isi kesimpulan tidak berupa pointpoint, namun berupa paragraf.

DAFTAR PUSTAKA

- Amit, S., Barua, L., & Kafy, A.-A. (2021). Countering violent extremism using social media and preventing implementable strategies for Bangladesh. *Heliyon*, 7(5), e07121. <https://doi.org/10.1016/j.heliyon.2021.e07121>
- Ananda Teresia and Fransiska Nangoy. (2025, January 14). Indonesia plans minimum age for social media use. *Reuters*. <https://www.reuters.com/world/asia-pacific/indonesia-planning-minimum-age-limit-social-media-users-minister-says-2025-01-14/>
- Andriani, A. D., Fitri, S. A., & Muchtar, K. (2024). MODEL KOMUNIKASI LITERASI DIGITAL DALAM MENGATASI UJARAN KEBENCIAN DI MEDIA SOSIAL. *Interaksi: Jurnal Ilmu Komunikasi*, 13(2), 439–464. <https://doi.org/10.14710/interaksi.13.2.439-464>
- Bulya, B., & Izzati, S. (2024). Indonesia’s Digital Literacy as a Challenge for Democracy in the Digital Age. *The Journal of Society and Media*, 8(2), 640–661. <https://doi.org/10.26740/jsm.v8n2.p640-661>
- Ghenai, A., Noorian, Z., Moradisani, H., Abadeh, P., Erentzen, C., & Zarrinkalam, F. (2025). Exploring hate speech dynamics: The emotional, linguistic, and thematic impact on social media users. *Information Processing & Management*, 62(3), 104079. <https://doi.org/10.1016/j.ipm.2025.104079>
- Giraud, E. H., Poole, E., de Quincey, E., & Richardson, J. E. (2025). Learning from online hate speech and digital racism: From automated to diffractive methods in social media analysis. *The Sociological Review*, 73(6), 1388–1407. <https://doi.org/10.1177/00380261241305260>
- Gutiérrez-Ángel, N., Sánchez-García, J.-N., Mercader-Rubio, I., García-Martín, J., & Brito-Costa, S. (2022). Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.896800>

- Hangartner, D., Gennaro, G., Alasiri, S., Bahrich, N., Bornhoff, A., Boucher, J., Demirci, B. B., Derksen, L., Hall, A., Jochum, M., Munoz, M. M., Richter, M., Vogel, F., Wittwer, S., Wüthrich, F., Gilardi, F., & Donnay, K. (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences*, 118(50), e2116310118. <https://doi.org/10.1073/pnas.2116310118>
- Harmawati, Y., Sapriya, Abdulkarim, A., Bestari, P., & Sari, B. I. (2024). Data of digital literacy level measurement of Indonesian students: Based on the components of ability to use media, advanced use of digital media, managing digital learning platforms, and ethics and safety in the use of digital media. *Data in Brief*, 54, 110397. <https://doi.org/10.1016/j.dib.2024.110397>
- Hassoun, A., Beacock, I., Carmody, T., Kelley, P. G., Goldberg, B., Kumar, D., Murray, L., Park, R. S., Sarmadi, B., & Consolvo, S. (2025). Beyond Digital Literacy: Building Youth Digital Resilience Through Existing “Information Sensibility” Practices. *Social Sciences*, 14(4), 230. <https://doi.org/10.3390/socsci14040230>
- Hirblinger, A. T. (2024). Fighting Hate Speech on Social Media: The Challenge of “Deep Work” in Digital Peacebuilding. In A. Reuss & S. Stetter (Eds.), *Social Media and Peacebuilding: How Digital Spaces Shape Conflict and Peace* (pp. 99–126). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-73917-0_5
- Ibrohim, M. O., & Budi, I. (2023). Hate speech and abusive language detection in Indonesian social media: Progress and challenges. *Heliyon*, 9(8), e18647. <https://doi.org/10.1016/j.heliyon.2023.e18647>
- Jia, Y., & Schumann, S. (2025). Tackling hate speech online: The effect of counter-speech on subsequent bystander behavioral intentions. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 19(1). <https://doi.org/10.5817/CP2025-1-4>
- Mansyur, M. S. (2020). *Model Literasi Digital untuk Melawan Ujaran Kebencian di Media Sosial (Digital Literacy Model to Counter Hate Speech on Social Media)*. <https://doi.org/10.33164/IPTEKKOM.22.2.2020.125-142>
- Minko, A. E. (2025, September 3). Building Digital Trust: Youth-Led Tech Solutions to Prevent Extremism in the Horn of Africa. *GNET*. <https://gnet-research.org/2025/09/03/building-digital-trust-youth-led-tech-solutions-to-prevent-extremism-in-the-horn-of-africa/>
- Obermaier, M., Schmid, U. K., & Rieger, D. (2025). Empowerment Is Key? How Perceived Political and Critical Digital Media Literacy Explain Direct and Indirect Bystander Intervention in Online Hate Speech. *Social Media + Society*, 11(1), 20563051251325598. <https://doi.org/10.1177/20563051251325598>
- Pukallus, S., & Arthur, C. (2024). Combating Hate Speech on Social Media: Applying Targeted Regulation, Developing Civil-Communicative Skills and Utilising Local Evidence-Based Anti-Hate Speech Interventions. *Journalism and Media*, 5(2), 467–484. <https://doi.org/10.3390/journalmedia5020031>

- Rees, J., & Montasari, R. (2023). The Use of Counter Narratives to Combat Violent Extremism Online. In R. Montasari, V. Carpenter, & A. J. Masy (Eds.), *Digital Transformation in Policing: The Promise, Perils and Solutions* (pp. 15–31). Springer International Publishing. https://doi.org/10.1007/978-3-031-09691-4_2
- Treacy, S., Reed, A., & Glazzard, A. (2024). A systematic integrative review of counter-messaging communication campaigns targeting terrorism or violent extremism. *Behavioral Sciences of Terrorism and Political Aggression*, 1–39. <https://doi.org/10.1080/19434472.2024.2419390>
- Veronika, R., Camelia, C., Febriliana, R., & Yapen, Y. (2023). DIGITAL LITERACY AS A SOCIAL MOBILIZATION AND LEARNING PLATFORM. *SEIKAT: Jurnal Ilmu Sosial, Politik Dan Hukum*, 2, 228–241. <https://doi.org/10.55681/seikat.v2i3.544>
- Wachs, S., Castellanos, M., Wettstein, A., Bilz, L., & Gámez-Guadix, M. (2023). Associations Between Classroom Climate, Empathy, Self-Efficacy, and Countering Hate Speech Among Adolescents: A Multilevel Mediation Analysis. *Journal of Interpersonal Violence*, 38(5–6), 5067–5091. <https://doi.org/10.1177/08862605221120905>
- Wachs, S., Wright, M. F., & Gámez-Guadix, M. (2024). From hate speech to HateLess. The effectiveness of a prevention program on adolescents' online hate speech involvement. *Computers in Human Behavior*, 157, 108250. <https://doi.org/10.1016/j.chb.2024.108250>
- Yang, F., Yao, R., Ren, Y., & Guo, L. (2025). Harmony in diversity: Digital literacy research in a multidisciplinary landscape. *Computers & Education*, 230, 105265. <https://doi.org/10.1016/j.compedu.2025.105265>
- Zahirah, A. (2025). The Role of Citizens in Countering Disinformation and Hate Speech on Social Media. *CIVITAS (JURNAL PEMBELAJARAN DAN ILMU CIVIC)*, 11(2), 19–31. <https://doi.org/10.36987/civitas.v11i2.7353>