
FAKE NEWS
HOAX
LIVE
Deepfake, Disinformasi, dan Krisis Kepercayaan Publik:
Strategi Literasi Media dan Kebijakan Regulasi
Oleh Rudy C Tarumingkeng
DEEPFAKE
BREAKING NEWS
BREAKING NEWS
1 DISINFORMASI

Oleh:
Prof Ir Rudy C Tarumingkeng, PhD
Professor of Management NUP: 9903252922
Rektor, Universitas Cenderawasih, Papua (1978-1988, dan
Rektor, Kampus AGRO Manokwari sekarang Universitas Papua Manokwari)
Coordinator, CIDA/DIKTI SFU Burnaby BC Canada 1988-1991
Rektor, Universitas Kristen Krida Wacana, Jakarta (1991-2000)
Ketua Dewan Guru Besar, IPB-University, Bogor (2005-2006)
AI - Data Analyst, dan Ketua Senat Akademik, IBM-ASMI, Jakarta 2024-

---

rudyct75@gmail.com
13 Februari 2026
Bogor, Indonesia

DEEPFAKE, DISINFORMASI, DAN KRISIS KEPERCAYAAN PUBLIK: STRATEGI LITERASI MEDIA DAN KEBIJAKAN REGULASI

Pendahuluan: ketika "bukti visual" tidak lagi menjadi bukti

Selama puluhan tahun, masyarakat modern membangun kebiasaan epistemik sederhana: *melihat adalah percaya*. Foto, rekaman suara, dan video dianggap "bukti paling kuat"—setidaknya lebih kuat dibanding rumor lisan atau teks anonim. Namun, kemunculan **deepfake** (media sintetis yang meniru wajah/gerak/ucapan manusia) dan industri **disinformasi** (informasi salah yang dibuat dan disebarkan secara sengaja untuk menipu) mengguncang fondasi kebiasaan itu. Perubahan ini bukan sekadar persoalan teknologi; ia adalah persoalan **tata kelola kepercayaan** (governance of trust) di ruang publik, termasuk kepercayaan pada pemilu, lembaga pemerintah, media, sains, dan bahkan pada relasi antarmanusia.

Di dalam ekosistem komunikasi digital, reputasi dan legitimasi sosial dibentuk oleh arus informasi yang sangat cepat: platform media sosial, aplikasi pesan, mesin rekomendasi, dan ekosistem *influencer* yang hidup

dari perhatian (attention economy). Ketika deepfake menjadi murah, mudah, dan realistis, disinformasi memperoleh "mesin produksi" baru: narasi palsu yang tidak hanya *dibaca*, tetapi juga *ditonton* dan *didengar* seperti kenyataan. Konsekuensinya adalah **krisis kepercayaan publik**: publik ragu pada informasi yang benar (karena takut tertipu), dan di saat yang sama mudah mempercayai informasi yang salah (karena kemasannya sangat meyakinkan). Fenomena ini sering disebut sebagai *liar's dividend*: ketika pemalsuan begitu mudah, pihak yang benar pun bisa dituduh palsu, sementara pelaku kebohongan berlindung di balik kabut keraguan.

Tulisan ini membahas secara akademik tiga lapis persoalan: (1) bagaimana deepfake dan disinformasi bekerja sebagai sistem; (2) mengapa ia merusak kepercayaan publik; dan (3) strategi *dua kaki* untuk merespons: **literasi media** (pendidikan-kognitif-sosial) dan **kebijakan regulasi** (norma hukum, kewajiban platform, serta mekanisme penegakan). Kerangka ini sengaja disusun bukan untuk "menang melawan hoaks sekali dua kali", melainkan untuk membangun **ketahanan epistemik** (epistemic resilience)—kemampuan masyarakat untuk tetap waras, kritis, dan adil dalam menilai informasi, di tengah banjir konten sintetis.

---

1) Konsep kunci: deepfake, disinformasi, dan "information disorder"

1.1 Deepfake sebagai media sintetis yang meniru realitas

Secara teknis, deepfake adalah konten gambar/audio/video yang dibuat atau dimanipulasi oleh AI sehingga menyerupai orang, objek, tempat, entitas, atau peristiwa nyata, dan tampak autentik padahal palsu. Regulasi **Uni Eropa** melalui *AI Act* memberi definisi "deep fake" sebagai "AI-generated or manipulated image, audio or video content… that would falsely appear… to be authentic or truthful." (EUR-Lex) Definisi ini

penting karena ia menegaskan dua unsur: (a) *keserupaan* dengan realitas, dan (b) *efek menipu* (false appearance of authenticity).

Kemampuan deepfake bertumpu pada kemajuan model generatif (seperti GAN dan model difusi), *voice cloning*, serta *face reenactment*. Kini seseorang dapat meniru suara publik figur hanya dari sampel audio pendek, atau menempelkan wajah pada tubuh lain di video, lalu menyelaraskan gerak bibir (lip-sync) dengan tingkat realisme tinggi. Jika dulu pemalsuan video membutuhkan studio produksi, hari ini ia bisa diproduksi oleh individu dengan perangkat konsumen.

1.2 Disinformasi, misinformasi, malinformasi

Untuk memahami dampak sosial-politik, kita perlu membedakan:

Misinformasi: informasi salah yang disebarkan tanpa niat menipu (misalnya seseorang meneruskan kabar palsu karena percaya).

Disinformasi: informasi salah yang sengaja dibuat/diorkestrasi untuk menipu, memanipulasi, atau meraih keuntungan politik/ekonomi.

Malinformasi: informasi benar yang disebarkan dengan konteks menyesatkan atau untuk mencelakai (misalnya doxing, potongan video benar tetapi diberi narasi palsu).

Kerangka "information disorder" yang banyak dirujuk menggarisbawahi bahwa masalah utama bukan sekadar "benar vs salah", tetapi **arsitektur produksi–distribusi–konsumsi** informasi yang memungkinkan manipulasi sistematis. (edoc.coe.int)

1.3 Kepercayaan publik sebagai "modal sosial" institusional

Kepercayaan publik bukan perasaan abstrak; ia adalah *modal sosial* yang membuat institusi dapat bekerja: warga patuh pada kebijakan karena percaya prosedurnya adil; investor percaya laporan karena percaya audit; pemilih percaya hasil karena percaya proses. Ketika deepfake menyerang, yang hancur bukan hanya citra individu, tetapi juga

mekanisme verifikasi sosial yang membuat masyarakat bisa mencapai kesepakatan minimal tentang kenyataan.

---

2) Mengapa deepfake mempercepat disinformasi: "mesin produksi" + "mesin distribusi"

2.1 Deepfake menurunkan biaya kebohongan

Dalam ekonomi politik informasi, *biaya produksi* menentukan skala. Deepfake menurunkan biaya kebohongan dengan tiga cara:

Otomatisasi: AI mengerjakan pekerjaan kreatif (editing, sinkronisasi bibir, pencahayaan, suara).

Personalisasi: konten dapat disesuaikan untuk target mikro (microtargeting) — misalnya variasi bahasa daerah, gaya bicara, atau referensi komunitas tertentu.

Plausibilitas: "bukti audiovisual" meningkatkan daya persuasi dibanding teks biasa, terutama untuk audiens yang tidak punya waktu memverifikasi.

Akibatnya, disinformasi tidak lagi tergantung pada jaringan propaganda besar; ia dapat dilakukan oleh aktor kecil dengan dampak besar.

2.2 Platform sebagai "mesin distribusi"

Konten tidak menyebar terutama karena "benar", tetapi karena *menarik*—memicu emosi, konflik identitas, atau rasa takut. Mesin rekomendasi platform cenderung memperkuat konten yang memicu keterlibatan (engagement). Di sini, deepfake menjadi "bahan bakar" yang sempurna: dramatis, memancing kemarahan, dan tampak nyata.

Karena itu, banyak kebijakan modern memaksa platform besar melakukan **penilaian risiko sistemik**. Di bawah *Digital Services Act* (DSA), penyedia *very large online platforms/search engines* wajib mengidentifikasi dan menilai risiko sistemik yang timbul dari desain

layanan dan sistem algoritmiknya. (eu-digital-services-act.com) Artinya, regulasi mulai menganggap disinformasi bukan sekadar "kesalahan pengguna", melainkan sebagian akibat desain ekosistem.

2.3 Disinformasi sebagai rantai pasok (supply chain)

Salah satu cara paling produktif memahami disinformasi adalah memetakannya seperti rantai pasok:

Sumber: aktor politik, kelompok kepentingan, *scammer*, atau "pabrik konten".

Produksi: generator deepfake, skrip narasi, akun palsu, bot.

Distribusi: platform, grup pesan, kanal video, iklan bertarget.

Amplifikasi: influencer, media partisan, akun jaringan.

Konsumsi: publik, komunitas, institusi.

Monetisasi/tujuan: uang (penipuan, iklan), kekuasaan (delegitimasi lawan), chaos (melemahkan institusi).

Melihatnya sebagai rantai pasok membantu kebijakan: intervensi dapat ditempatkan di setiap mata rantai, bukan hanya di "ujung" (menghapus konten setelah viral).

---

3) Dampak sosial: dari penipuan finansial sampai delegitimasi demokrasi

3.1 Keamanan dan konflik: contoh deepfake dalam perang informasi

Deepfake digunakan untuk operasi psikologis: menurunkan moral, memecah belah, atau menimbulkan kebingungan. Contoh yang sering dikutip adalah video deepfake yang menampilkan Presiden **Volodymyr Zelenskyy** seolah menyerukan pasukan Ukraina untuk menyerah (2022). (Reuters) Meskipun cepat dibantah, episode ini menunjukkan bagaimana deepfake dipakai sebagai instrumen perang informasi: bukan untuk

"membuat orang percaya selamanya", tetapi cukup untuk menciptakan keraguan sesaat yang merusak koordinasi dan kepercayaan.

3.2 Demokrasi dan pemilu: "kampanye berbasis ilusi"

Dalam konteks pemilu, deepfake bisa:

memfitnah kandidat (scandal palsu),

memalsukan pernyataan (policy stance palsu),

menciptakan *suppression* (menipu pemilih agar tidak datang),

atau mengacaukan hari pemungutan (informasi palsu tentang TPS, waktu, prosedur).

Kasus *robocall* yang meniru suara Presiden **Joe Biden** untuk memengaruhi pemilih dalam konteks pemilihan pendahuluan di New Hampshire menjadi contoh bagaimana suara sintetis dipakai untuk manipulasi politik. Di level kebijakan, **Federal Communications Commission** pada 2024 menyatakan bahwa suara AI dalam robocall termasuk "artificial/prerecorded voice" sehingga berada dalam rezim larangan/aturan *Telephone Consumer Protection Act*.

Pesan intinya: deepfake memindahkan manipulasi politik dari level retorika ke level "bukti" audiovisual—yang selama ini menjadi jantung persuasi politik modern.

3.3 Kriminalitas ekonomi: penipuan identitas, pemerasan, dan fraud korporat

Deepfake juga menguatkan *social engineering*: penipu meniru suara atasan untuk memerintahkan transfer dana, meniru wajah dalam video call, atau mengirim video "bukti" palsu. Dengan meningkatnya penggunaan verifikasi jarak jauh, deepfake membuat standar keamanan berbasis audio/video menjadi rapuh. Ini mendorong kebutuhan autentikasi tambahan: *liveness detection*, verifikasi multi-faktor, dan protokol *call-back*.

3.4 Kekerasan berbasis gender: deepfake intim tanpa persetujuan

Salah satu dampak paling destruktif adalah deepfake pornografi non-konsensual—sering menarget perempuan publik figur. Kasus deepfake terhadap **Taylor Swift** pada 2024 menunjukkan bagaimana media sintetis dapat menjadi alat pelecehan massal dan memperlihatkan keterlambatan platform serta regulasi dalam melindungi korban. (Wikipedia)

Sebagai respons kebijakan, **Britania Raya** mempercepat langkah kriminalisasi pembuatan atau permintaan deepfake intim tanpa persetujuan. (GOV.UK) Ini menandai pergeseran: deepfake tidak hanya dipandang sebagai isu "konten", tetapi isu *hak tubuh digital* (digital bodily integrity).

---

4) Mengapa ini memicu krisis kepercayaan publik: lima mekanisme sosiologis

4.1 Erosi "otoritas epistemik"

Dalam masyarakat modern, kebenaran publik ditopang oleh institusi: media profesional, lembaga ilmiah, pengadilan, otoritas pemilu. Disinformasi dan deepfake menyerang kredibilitas institusi ini dengan dua strategi:

Delegitimasi: "media bohong", "data palsu", "hasil pemilu dicurangi".

Overload: banjir klaim sehingga publik lelah memeriksa, lalu menyerah pada sinisme.

4.2 Polarisasi identitas

Disinformasi efektif ketika ia selaras dengan identitas kelompok. Deepfake memberi "alat bukti" untuk mengukuhkan prasangka. Masyarakat tidak lagi bertanya "apakah benar?", melainkan "apakah ini menguntungkan kelompok saya?".

4.3 Normalisasi kecurigaan: *everything can be fake*

Ketika publik tahu deepfake ada, reaksi psikologisnya sering bukan menjadi lebih kritis, tetapi menjadi lebih skeptis terhadap semuanya—termasuk informasi benar. Ini membunuh kemampuan masyarakat membangun konsensus minimal.

4.4 Krisis prosedural

Kepercayaan publik bukan hanya tentang hasil, tetapi tentang prosedur: transparansi, akuntabilitas, hak jawab, mekanisme koreksi. Disinformasi mengacaukan prosedur koreksi karena kecepatan penyebaran jauh lebih tinggi daripada kecepatan klarifikasi.

4.5 "Kesenjangan literasi" sebagai ketimpangan baru

Orang yang memiliki keterampilan verifikasi (lateral reading, cek sumber, cek metadata) menjadi lebih aman, sedangkan yang tidak memilikinya menjadi rentan. Krisis kepercayaan lalu berubah menjadi ketimpangan sosial: kelompok rentan paling mudah dimanipulasi.

---

5) Strategi literasi media: membangun "imunitas kognitif" masyarakat

Literasi media yang efektif tidak cukup berupa slogan "jangan mudah percaya". Ia harus diperlakukan seperti *kapasitas organisasi*: punya metode, latihan, indikator, dan institusi pendukung (sekolah, kampus, komunitas, media, pemerintah).

5.1 Literasi media sebagai kemampuan prosedural (bukan sekadar pengetahuan)

Banyak program literasi gagal karena terlalu fokus pada *pengetahuan deklaratif* ("hoaks itu berbahaya"), bukan *pengetahuan prosedural* ("bagaimana memeriksa klaim secara cepat"). Dua pendekatan yang kuat:

Lateral reading (membaca menyamping): ketika menemukan klaim, kita tidak "tenggelam" di satu situs, tetapi segera membuka sumber lain yang kredibel untuk memverifikasi. Penelitian pendidikan kewargaan digital oleh **Stanford History Education Group** menunjukkan bahwa strategi ini membedakan penilai informasi yang andal dari yang mudah tertipu. (Stanford Digital Repository)

Metode SIFT: Stop, Investigate the source, Find better coverage, Trace claims. Metode ini populer karena sederhana, cepat, dan sesuai ritme konsumsi informasi modern.

Dalam praktik, literasi media bukan mengubah semua warga menjadi "detektif", tetapi memberi *kebiasaan minimum* agar tidak menjadi korban manipulasi massal.

5.2 Literasi deepfake: indikator, kebiasaan verifikasi, dan batasannya

Pelatihan deepfake harus realistis: mengajarkan bahwa "mencari cacat visual" (mis. kedipan mata aneh) semakin tidak memadai. Fokus harus beralih ke:

verifikasi konteks: apakah peristiwa ini dilaporkan media kredibel lain?

jejak sumber: dari akun mana pertama kali muncul? apakah akun itu terpercaya?

pembuktian lintas kanal: adakah rekaman lain dari sudut berbeda?

forensik ringan: cek metadata jika tersedia, cek tanggal unggah, cek geolokasi (bila relevan).

Pada level warga, kita mengajarkan *prinsip kehati-hatian*: konten yang "terlalu dramatis" dan "terlalu pas dengan prasangka" harus diperlakukan sebagai *high-risk content*.

5.3 Prebunking dan inoculation: "vaksin" sebelum terpapar

Pendekatan **inoculation theory** berargumen: lebih efektif memberi "vaksin kognitif" sebelum hoaks datang, daripada memperbaiki setelah

viral. Eksperimen *prebunking* (memberi contoh teknik manipulasi seperti scapegoating, false dilemma, impersonation) menunjukkan efek protektif terhadap misinformasi. (prebunking.withgoogle.com)

Dalam konteks deepfake, prebunking bisa berupa:

mengenalkan taktik "impersonation" (meniru tokoh),

taktik "fabricated evidence" (bukti video palsu),

taktik "context collapse" (memotong video benar untuk narasi palsu).

5.4 Literasi sebagai ekologi: sekolah–media–komunitas–platform

Literasi media yang efektif harus bersifat ekosistem:

Sekolah & kampus: modul wajib literasi digital (verifikasi sumber, bias kognitif, etika berbagi).

Media profesional: transparansi proses verifikasi, rubrik "bagaimana kami memeriksa".

Komunitas: relawan cek fakta, kelas publik, komunitas anti-hoaks seperti **Mafindo**. (mafindo.or.id)

Platform: desain yang menambahkan "friction" (mis. peringatan sebelum berbagi konten yang dipertanyakan) dan mendukung pelaporan.

UNESCO telah lama mendorong *Media and Information Literacy* (MIL) sebagai kapasitas warga abad ke-21, menekankan dimensi kritis, etis, dan partisipatif. (UNESCO)

5.5 Narasi kasus (Indonesia): simulasi pelatihan untuk komunitas lokal

Bayangkan sebuah kota di Indonesia menjelang pilkada. Di grup WhatsApp RT/RW beredar video seorang kandidat "menghina" agama tertentu. Video itu realistis, emosi warga naik, dan dalam 3 jam suasana sosial memanas. Pelatihan literasi media yang baik mengajarkan prosedur 10 menit:

Stop: jangan sebar.

Investigate: cek akun sumber pertama—apakah akun resmi?

Find better coverage: cari berita di media arus utama, cek klarifikasi KPU/Bawaslu/korban.

Trace: apakah video punya versi lebih panjang? kapan pertama diunggah?

Decide: jika tidak terverifikasi, perlakukan sebagai manipulasi.

Di tahap lanjut, pelatihan juga membangun *norma sosial*: "tidak menyebarkan sebelum verifikasi" menjadi standar kehormatan komunitas—seperti etika antre atau etika berlalu lintas.

---

6) Teknologi penandaan dan autentikasi: dari watermark hingga provenance

Literasi saja tidak cukup, karena beban verifikasi tidak bisa sepenuhnya dipindahkan ke individu. Maka muncul agenda "infrastruktur kepercayaan" (trust infrastructure): penandaan konten AI, pelacakan asal konten, dan standar autentikasi.

6.1 Kewajiban penandaan dalam EU AI Act

EU AI Act memuat kewajiban penting:

Penyedia sistem AI yang menghasilkan konten sintetis (audio/gambar/video/teks) harus memastikan output *ditandai* dan *terdeteksi* sebagai buatan/manipulasi AI dalam format yang bisa dibaca mesin, sejauh layak secara teknis. (EUR-Lex)

Pengguna/deployer yang memakai AI untuk membuat deepfake wajib **mengungkapkan** bahwa konten dibuat atau dimanipulasi. Ada pengecualian tertentu (mis. penegakan hukum) dan perlakuan khusus untuk karya satir/artistik agar tidak mematikan ekspresi. (EUR-Lex)

Poin kebijakannya jelas: transparansi bukan sekadar etika, tetapi **kewajiban**.

6.2 Standar dan problem "watermarking yang rapuh"

Secara teknis, watermark dan metadata dapat dihapus, dirusak, atau hilang ketika konten diunggah ulang. Karena itu, pendekatan autentikasi perlu kombinasi: watermark + metadata + tanda kriptografis + kebijakan platform + audit. Kajian kebijakan juga mengingatkan bahwa mandat watermark tanpa standar interoperabilitas dan tanpa ekosistem penegakan dapat jatuh menjadi simbolik. (Data Innovation)

6.3 China: regulasi "deep synthesis" dan kewajiban label

Di Tiongkok, regulasi deep synthesis mulai berlaku pada 2023 dan mencakup kewajiban pelabelan konten yang dihasilkan/diedit dengan teknologi deep synthesis. Ringkasan **Library of Congress** menegaskan berlakunya ketentuan tersebut dan aktor regulator terkait. (The Library of Congress) Sementara teks terjemahan menunjukkan struktur aturan yang menempatkan pelabelan sebagai kewajiban penyedia layanan deep synthesis. (China Law Translate)

---

7) Kebijakan regulasi: prinsip, opsi desain, dan risiko kebijakan yang salah arah

Regulasi deepfake/disinformasi adalah medan sulit: jika terlalu lemah, manipulasi merajalela; jika terlalu keras, kebebasan berekspresi dan hak sipil bisa tertekan. Karena itu, kebijakan harus bertumpu pada prinsip *proportionality* dan *due process*.

7.1 Prinsip dasar regulasi yang sehat

Transparansi: kewajiban label/disclosure untuk konten sintetis berisiko tinggi.

Akuntabilitas: siapa pembuat, penyebar, dan platform harus punya tanggung jawab proporsional.

Perlindungan korban: jalur cepat *takedown* untuk deepfake intim, pemerasan, dan impersonation.

Due process: mekanisme banding agar tidak terjadi sensor sewenang-wenang.

Interoperabilitas: standar teknis yang kompatibel lintas platform.

Koordinasi lintas lembaga: karena disinformasi menyentuh pemilu, keamanan, perlindungan konsumen, data pribadi, dan pendidikan.

7.2 Opsi kebijakan (policy menu) yang umum dipakai negara

Kewajiban label untuk konten AI (provider & deployer).

Kewajiban platform melakukan *risk assessment* dan mitigasi (model DSA). (eu-digital-services-act.com)

Kriminalisasi spesifik untuk deepfake intim non-konsensual (model UK). (GOV.UK)

Penegakan sektor telekomunikasi untuk robocall suara AI (model FCC).

Kode etik/kode praktik multi-pihak (co-regulation), seperti *EU Code of Practice on Disinformation* yang diperkuat 2022 dan diarahkan menjadi *Code of Conduct* di bawah DSA. (Digital Strategy)

Kekuatan pendekatan co-regulation: adaptif terhadap teknologi. Kelemahannya: sering bergantung pada kepatuhan sukarela dan asimetri kekuatan antara regulator–platform.

7.3 Risiko kebijakan yang harus dihindari

Over-criminalization: semua "konten salah" dipidana → berisiko membungkam kritik dan jurnalisme investigatif.

Definisi kabur: "disinformasi" tanpa definisi operasional → rawan disalahgunakan.

Penegakan selektif: hukum dipakai untuk lawan politik → memperparah krisis kepercayaan.

Beban pada warga: jika regulasi hanya menyuruh "masyarakat harus cerdas" tanpa memaksa platform memperbaiki desain.

---

8) Konteks Indonesia: payung hukum yang relevan dan peluang desain kebijakan

Di Indonesia, respons terhadap disinformasi dan deepfake bertumpu pada beberapa pilar regulasi umum, meski belum selalu spesifik menyasar "konten sintetis".

8.1 UU PDP: perlindungan data dan fondasi kepercayaan digital

Undang-Undang No. 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP) membangun fondasi hak dan kewajiban dalam pemrosesan data pribadi. (BPK Regulations) Dalam konteks deepfake, UU PDP relevan karena deepfake sering memakai data biometrik (wajah, suara) yang dapat menjadi data sangat sensitif. Dengan demikian, tata kelola data (consent, security, breach handling) adalah bagian dari pencegahan deepfake.

8.2 UU ITE hasil perubahan 2024: ruang penataan ulang keseimbangan

Undang-Undang No. 1 Tahun 2024 (perubahan kedua UU ITE) memperbarui beberapa ketentuan yang selama ini memicu perdebatan publik dan multitafsir. (BPK Regulations) Dalam agenda deepfake, UU ITE relevan sebagai dasar penegakan atas konten merugikan, penghinaan, pemerasan, penipuan elektronik, dan kewajiban penyelenggara sistem elektronik—namun tantangan utamanya adalah memastikan *due process* dan perlindungan kebebasan berekspresi.

8.3 Pedoman etika AI (soft law) dan kebutuhan "hard law" yang presisi

Kementerian yang kini dikenal sebagai Kementerian Komunikasi dan Digital (sebelumnya Kominfo) menerbitkan pedoman etika AI melalui surat edaran (aturan lunak). Liputan tentang Surat Edaran Menkominfo No. 9 Tahun 2023 menekankan nilai seperti transparansi, akuntabilitas, keamanan, dan perlindungan data pribadi. (Green Network Asia - Indonesia) Pedoman ini penting sebagai *norm-setting*, tetapi untuk deepfake berisiko tinggi (pemilu, fraud, pornografi non-konsensual), dibutuhkan perangkat "hard law" yang lebih operasional: definisi, kewajiban label, sanksi, dan mekanisme penegakan.

---

9) Rekomendasi strategi terpadu: literasi + regulasi + desain platform

Agar tidak terjebak pada pendekatan reaktif (selalu mengejar hoaks yang sudah viral), respons harus berbentuk strategi terpadu dan berlapis.

9.1 Arsitektur kebijakan "3L": Literacy, Labeling, Liability

Literacy:

Kurikulum literasi media berbasis prosedur (SIFT + lateral reading). (Stanford Digital Repository)

Program prebunking untuk teknik manipulasi (impersonation, fabricated evidence). (prebunking.withgoogle.com)

Sertifikasi pelatihan untuk aparat, humas pemerintah, jurnalis lokal, guru.

Labeling:

Kewajiban label/disclosure untuk konten sintetis berisiko tinggi (audio/video).

Dorong standar teknis penandaan (machine-readable) dan interoperabilitas, belajar dari EU AI Act. (EUR-Lex)

Liability:

Tanggung jawab platform berbasis risiko (risk assessment dan mitigasi). (eu-digital-services-act.com)

Jalur cepat perlindungan korban (deepfake intim, pemerasan, impersonation). (GOV.UK)

Penegakan sektoral untuk penipuan telekomunikasi/robocall (pelajaran dari FCC).

9.2 Roadmap implementasi 18 bulan (contoh desain kebijakan operasional)

(A) 0–6 bulan: kesiapan institusional

Bentuk *task force* lintas lembaga (komunikasi digital, penegak hukum, pemilu, perlindungan konsumen, pendidikan).

Standarkan SOP respons cepat: verifikasi, klarifikasi publik, kerja sama platform.

Bangun *incident database* (anonim) untuk pola deepfake & disinformasi.

(B) 6–12 bulan: kapasitas publik dan standar teknis

Modul literasi media nasional: SIFT + lateral reading + prebunking. (Stanford Digital Repository)

Uji coba label konten AI pada beberapa platform/instansi.

Pelatihan jurnalis dan humas pemerintah: protokol "klarifikasi berbasis bukti" (hindari pernyataan emosional yang memperbesar viralitas).

(C) 12–18 bulan: regulasi presisi dan audit platform

Rumuskan aturan turunan/specific law untuk deepfake berisiko tinggi:

definisi operasional,

kewajiban disclosure,

perlindungan korban (takedown cepat),

sanksi proporsional,

mekanisme banding.

Audit risiko platform besar dengan indikator: waktu respons, tingkat penghapusan konten berbahaya, transparansi label.

9.3 Indikator keberhasilan (karena tanpa metrik, strategi mudah jadi retorika)

Time-to-clarify: berapa jam rata-rata klarifikasi resmi muncul setelah konten viral.

Time-to-action: berapa jam rata-rata konten berbahaya diturunkan setelah laporan valid.

Prevalensi: proporsi konten sintetis tak berlabel pada topik-topik berisiko.

Trust index: survei kepercayaan pada institusi pemilu, media, dan kanal klarifikasi resmi.

Literacy gain: peningkatan skor uji literasi (pretest–posttest) pada siswa/komunitas.

---

10) Penutup: membangun kembali kepercayaan di era "realitas yang dapat diproduksi"

Deepfake dan disinformasi bukan badai sesaat, melainkan perubahan struktural pada cara masyarakat memproduksi dan menguji kebenaran. Karena itu, respons yang paling masuk akal adalah membangun **rezim kepercayaan baru**: literasi media sebagai kebiasaan warga, regulasi sebagai pagar keadilan, dan desain platform sebagai infrastruktur pencegahan.

Pelajaran dari berbagai rezim kebijakan menunjukkan arah yang konsisten: transparansi (label/disclosure) menjadi norma, platform

dipaksa mengelola risiko sistemik, korban mendapat perlindungan cepat, dan pendidikan publik diarahkan pada keterampilan verifikasi yang praktis. EU AI Act menegaskan kewajiban penandaan dan disclosure untuk konten sintetis. (EUR-Lex) DSA menempatkan risiko disinformasi sebagai risiko sistemik yang harus dinilai dan dimitigasi oleh platform besar. (eu-digital-services-act.com) China menempuh jalur pelabelan dan tata kelola layanan deep synthesis. (The Library of Congress) Sementara penegak telekomunikasi di AS bergerak menutup celah robocall suara AI.

Bagi Indonesia, tantangannya adalah menyusun kebijakan yang presisi: cukup kuat untuk melindungi warga dan institusi demokrasi, tetapi cukup adil untuk menjaga kebebasan berekspresi serta mencegah penegakan yang sewenang-wenang. Dengan fondasi UU PDP dan pembaruan UU ITE, serta pedoman etika AI, ruang perbaikan itu terbuka—tinggal memastikan integrasi, standar teknis, dan kapasitas pelaksanaannya. (BPK Regulations)

Pada akhirnya, krisis kepercayaan publik tidak dapat diselesaikan hanya dengan "melawan hoaks", tetapi dengan **membangun ulang ekosistem kebenaran**—sebuah ekosistem yang memungkinkan warga untuk ragu secara sehat, memeriksa dengan cepat, dan percaya kembali secara rasional.

Berikut **Glosarium** dan **Referensi** untuk artikel Anda. (Saya juga sertakan file cover yang sudah dibuat.)

Cover (1 halaman):
Download cover PNG

Glosarium

AI-generated content (Konten buatan AI): Konten (teks, gambar, audio, video) yang dihasilkan oleh sistem AI, bukan direkam langsung dari peristiwa nyata.

Algorithmic amplification (Amplifikasi algoritmik): Penguatan penyebaran konten oleh sistem rekomendasi platform berdasarkan metrik keterlibatan (klik, komentar, durasi tonton).

Artificial voice / voice cloning (Suara sintetis/kloning suara): Replikasi karakter suara seseorang menggunakan model AI, sering dipakai untuk penipuan dan impersonasi.

Astroturfing: Penciptaan kesan seolah dukungan publik "organik", padahal dikendalikan terkoordinasi (akun palsu, bot, jaringan).

Attention economy (Ekonomi perhatian): Ekosistem digital yang memonetisasi perhatian pengguna; konten yang memancing emosi cenderung menang.

Authenticity (Keautentikan): Tingkat kesesuaian konten dengan kejadian nyata dan sumber asli yang dapat diverifikasi.

Bot: Akun otomatis yang diprogram untuk memposting, menyukai, mengikuti, atau meramaikan percakapan agar narasi tertentu tampak populer.

Content moderation (Moderasi konten): Kebijakan dan tindakan platform untuk menilai, membatasi, memberi label, atau menghapus konten melanggar aturan/hukum.

Debunking: Koreksi terhadap klaim salah setelah beredar (klarifikasi, cek fakta, penjelasan konteks).

Deepfake: Konten sintetis (gambar/audio/video) yang dimanipulasi agar tampak autentik—misalnya wajah/ucapan seseorang ditiru.

Deep synthesis: Istilah regulasi (mis. China) untuk layanan yang menggunakan teknologi generatif/manipulatif tingkat lanjut, termasuk deepfake.

Disinformation (Disinformasi): Informasi salah yang dibuat dan disebarkan **secara sengaja** untuk menipu/manipulasi atau memperoleh keuntungan.

Doxing: Penyebaran data pribadi sensitif (alamat, nomor, identitas) untuk mengintimidasi atau mencelakai.

Election integrity (Integritas pemilu): Keutuhan proses pemilu yang bebas manipulasi, transparan, dapat diaudit, dan dipercaya publik.

Friction (Gesekan berbagi): Intervensi desain (mis. peringatan, jeda, prompt verifikasi) agar pengguna tidak impulsif menyebarkan konten.

Information disorder: Kerangka yang memetakan misinformasi, disinformasi, dan malinformasi sebagai polusi informasi sistemik.

Lateral reading: Teknik evaluasi sumber dengan "membaca menyamping"—membandingkan klaim lewat sumber lain yang lebih kredibel.

Liar's dividend: Situasi ketika keberadaan deepfake membuat pihak bersalah bisa menyangkal bukti ("itu deepfake"), dan publik makin bingung.

Malinformation (Malinformasi): Informasi benar yang disebarkan dengan konteks menyesatkan atau tujuan merugikan (mis. potongan video tanpa konteks).

Misinformation (Misinformasi): Informasi salah yang disebarkan tanpa niat menipu (biasanya karena percaya).

Microtargeting: Penargetan pesan sangat spesifik pada segmen kecil (wilayah, minat, emosi, identitas) untuk memaksimalkan efek persuasi.

Provenance (Asal-usul konten): Riwayat pembuatan/penyuntingan/unggahan konten; penting untuk verifikasi.

Robocall / robotext: Panggilan/pesan otomatis berskala besar; berbahaya bila memakai suara AI untuk penipuan/disinformasi.

Risk assessment (Penilaian risiko sistemik): Kewajiban menilai risiko yang timbul dari desain layanan/platform terhadap masyarakat (mis. disinformasi).

SIFT: Metode cepat evaluasi informasi: *Stop–Investigate the source–Find better coverage–Trace to original context*.

Synthetic media (Media sintetis): Payung istilah untuk konten yang dihasilkan/dimodifikasi AI sehingga menyerupai realitas.

Transparency obligation (Kewajiban transparansi): Kewajiban melabeli/menyatakan konten AI agar pengguna tahu itu buatan/manipulasi.

Verification (Verifikasi): Proses pembuktian klaim dengan sumber primer/sekunder yang kredibel dan dapat diaudit.

Watermarking (Watermark digital): Penandaan konten (terlihat/tidak terlihat) untuk menunjukkan konten buatan AI atau untuk pelacakan.

UU PDP: Undang-Undang Pelindungan Data Pribadi; relevan untuk penyalahgunaan data biometrik (wajah/suara) dalam deepfake.

UU ITE (Perubahan 2024): Payung hukum transaksi elektronik/informasi; dapat terkait penegakan terhadap konten merugikan dan pelanggaran elektronik.

Referensi (pilihan inti)

Council of Europe. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. (edoc.coe.int)

Claire Wardle, & Hossein Derakhshan. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. (Laporan kerangka mis-/dis-/malinformation). (Council of Europe)

UNESCO. (n.d.). *Media and Information Literacy (MIL)*. (UNESCO)

Mike Caulfield. (2019). *SIFT: The four moves* (kerangka evaluasi cepat informasi online). (Hapgood)

Stanford History Education Group; Joel Breakstone et al. (2019). *Students' civic online reasoning: A national portrait*. (Stanford Digital Repository)

European Union. (2022). *Regulation (EU) 2022/2065 (Digital Services Act) — Article 34 (Risk assessment)*. (eu-digital-services-act.com)

—. (2024). *AI Act — Article 50 (Transparency obligations; marking/labeling AI-generated or manipulated content, termasuk deepfake)*. (Artificial Intelligence Act)

Federal Communications Commission. (2024, February 8). *FCC makes AI-generated voices in robocalls illegal* (penegasan suara AI termasuk "artificial voice" dalam rezim TCPA). (Federal Communications Commission)

Government of the United Kingdom. (2025, January 7). *Government crackdown on explicit deepfakes* (arah kriminalisasi/penindakan deepfake seksual). (GOV.UK)

University of Cambridge, BBC Media Action, & Jigsaw. (2022/2023). *A practical guide to prebunking misinformation* (panduan intervensi "inoculation/prebunking"). (prebunking.withgoogle.com)

China Law Translate. (2022). *Provisions on the administration of deep synthesis internet information services* (terjemahan aturan "deep synthesis"). (China Law Translate)

Republik Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP)*. (BPK Regulations)

—. (2024). *Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas UU Nomor 11 Tahun 2008 (UU ITE)*. (BPK Regulations)

Organisation for Economic Co-operation and Development. (2024). *Good practice principles for public communication responses to mis- and disinformation* (panduan komunikasi publik menghadapi mis-/disinformasi). (OECD)

(Opsional pengayaan kebijakan) *European Commission press release on integrating the Code of Practice on Disinformation into the DSA framework* (2025). (European Commission)

Copilot for this article - Chatgpt 5.2 Thinking. Access date: 13 Februari 2026. Prompting on Writer's account (Rudy C Tarumingkeng) https://chatgpt.com/c/698e7ccd-37a8-839e-b6c7-abcf1a7e7567