

# Model EPAM-2025 Untuk Analisis Keselarasan Opini Publik dan Kebijakan Literasi Digital

Tri Wiyono\*, Muhammad Irfan Sarif, Andika Dwi Aryo H, Ahmad Syaukani, Muhammad Zikri Ramadhan

Pascasarjana, Magister Teknologi Informasi, Universitas Pembangunan Pancabudi, Medan, Indonesia

Email: <sup>1</sup>\*tri82.wiyono@gmail.com, <sup>2</sup>irfanberbagi@gmail.com, <sup>4</sup>kaniharahap@gmail.com,

<sup>5</sup>muhammadzikriramadhan7@gmail.com

Email Penulis Korespondensi: tri82.wiyono@gmail.com\*

Submitted: 22/11/2025; Accepted: 09/12/2025; Published: 31/12/2025

**Abstrak**— Transformasi digital yang berlangsung cepat menuntut peningkatan literasi digital yang lebih komprehensif. Namun, berbagai indikator nasional menunjukkan penurunan pada aspek etika digital, keamanan siber, dan kemampuan berpikir kritis. Penelitian ini bertujuan menganalisis kesenjangan antara opini publik dan kebijakan literasi digital nasional melalui pendekatan komputasional berbasis *Natural Language Processing* (NLP). Untuk tujuan tersebut, dikembangkan model hibrida orisinal EPAM-2025 (*Entity–Policy Alignment Model*) sebagai kerangka pengukuran keselarasan kebijakan. Dataset penelitian terdiri atas 2.165 tweet berbahasa Indonesia yang diperoleh secara etis melalui API platform X (Twitter). Prosedur analisis mencakup pembersihan data, tokenisasi, ekstraksi entitas menggunakan Named Entity Recognition (NER), analisis sentimen, serta pengukuran kesamaan semantik berbasis *TF-IDF cosine similarity*. Skor keselarasan dihitung menggunakan formula  $S_a = \alpha S_m + \beta S_s(\text{norm})$ . Hasil penelitian menunjukkan dominasi opini netral (92,24%) serta tingkat kesamaan semantik yang sangat rendah ( $<0,1$ ), menandakan bahwa terminologi kebijakan digital seperti *digital safety* dan *digital ethics* belum terinternalisasi dalam wacana publik. Model EPAM-2025 juga menunjukkan performa evaluatif yang stabil dengan klasifikasi tepat pada dua kategori aktif (“Tidak Selaras” dan “Perlu Analisis Lanjut”). Penelitian ini memberikan kontribusi metodologis melalui pengembangan pendekatan kuantitatif yang objektif untuk mengukur keselarasan opini publik terhadap kebijakan, serta membuka peluang pemanfaatan analitik opini publik dalam mendukung perumusan kebijakan nasional berbasis bukti.

**Kata Kunci:** Literasi Digital; NLP; NER; EPAM-2025; X; Twitter; Opini Publik

**Abstract**— The rapid acceleration of digital transformation necessitates more comprehensive enhancement of digital literacy. However, national indicators show declining performance in digital ethics, cybersecurity awareness, and critical thinking skills. This study aims to analyze the gap between public opinion and national digital literacy policy using a computational approach based on Natural Language Processing (NLP). To achieve this, the study introduces EPAM-2025 (Entity–Policy Alignment Model), a hybrid analytical framework for measuring policy alignment. The dataset comprises 2,165 Indonesian-language tweets ethically collected through the X (Twitter) API. The analytical procedures include data cleaning, tokenization, entity extraction using Named Entity Recognition (NER), sentiment analysis, and semantic similarity measurement using TF-IDF cosine similarity. The final alignment score is computed using  $S_a = \alpha S_m + \beta S_s(\text{norm})$ . The findings reveal a dominance of neutral opinions (92.24%) and very low semantic similarity scores ( $<0.1$ ), indicating that policy terminologies such as *digital safety* and *digital ethics* are not yet internalized within public discourse. EPAM-2025 also demonstrates stable evaluative performance, correctly classifying two active categories (“Unaligned” and “Requires Further Analysis”). This study contributes methodologically by offering an objective quantitative approach to measuring public–policy alignment and highlights opportunities for utilizing public opinion analytics to support evidence-based national digital literacy policy development.

**Keywords:** Digital Literacy; NLP; NER; EPAM-2025; X; Twitter, Public Opinion

## 1. PENDAHULUAN

Ditengah percepatan transformasi digital yang semakin intensif, literasi digital menjadi kompetensi dasar yang tidak dapat dipisahkan dari kehidupan warga negara. Kemampuan untuk mengakses, mengevaluasi, dan memanfaatkan informasi digital secara efektif kini dipandang sebagai prasyarat bagi partisipasi sosial yang bertanggung jawab [1]. Meskipun penetrasi internet dan penggunaan perangkat digital di Indonesia terus menunjukkan peningkatan, sejumlah penelitian menegaskan bahwa tingkat literasi digital masyarakat masih menghadapi tantangan yang cukup besar [2], [3]. Tantangan tersebut tidak hanya menyangkut keterampilan teknis, tetapi juga mencakup kemampuan berpikir kritis dan sensitivitas etis dalam merespons berbagai arus informasi di ruang digital [4], sehingga muncul pertanyaan utama sejauh mana opini publik selaras dengan kebijakan literasi digital nasional.

Penelitian-penelitian terkini memberikan gambaran yang lebih komprehensif mengenai persoalan tersebut. Paramitha (2023) [1] menegaskan bahwa rendahnya literasi digital berdampak langsung pada kapasitas masyarakat dalam memahami risiko digital dan mempraktikkan etika bermedia [5]. Afrina (2024) [2], melalui analisis konten media daring, menemukan bahwa sebagai pengguna internet di Indonesia masih berada pada kategori literasi digital rendah [6], terutama pada aspek evaluasi informasi. Smith dan Storrs (2023) [3] menunjukkan bahwa literasi digital pada kelompok muda cenderung berfokus pada aspek teknis, sementara kemampuan berpikir kritis dan pertimbangan etis masih lemah. Temuan serupa diungkap oleh Putrayasa (2024) [4] yang menyimpulkan bahwa tantangan utama literasi digital di Indonesia terletak pada lemahnya kemampuan reflektif dan etika digital generasi muda. Lebih jauh Bulya dan Izzati (2024) [7] menekankan bahwa rendahnya literasi digital mengancam kualitas

demokrasi digital karena memengaruhi pola komunikasi publik dan kerentanan masyarakat terhadap disinformasi [8].

Di sisi lain, beberapa kajian telah menempatkan media sosial sebagai arena penting pembentukan opini digital [9]. Bulya dan Izzati (2024) [7] menunjukkan bahwa interaksi sosial di platform digital berperan signifikan dalam membentuk opini politik dan kebijakan. Hermanto et al. (2024) [10] menegaskan bahwa X (Twitter) merupakan ruang diskursif yang sangat dinamis dan merefleksikan persepsi publik secara real-time. Meta-analisis NLP oleh Jim et al. (2024) [11] menunjukkan bahwa metode NLP efektif untuk memetakan pola sentimen dan dinamika opini publik. Selain itu, penelitian Wang et al. (2023) [12] membuktikan bahwa *Named Entity Recognition* (NER) mampu mengidentifikasi aktor dan isu kunci secara akurat dalam wacana publik digital. Hidayat et al. (2020) [13] juga menemukan bahwa analisis wacana digital memberikan gambaran awal mengenai bagaimana masyarakat memahami dan merespons kebijakan pemerintah di ruang daring.

Meskipun demikian, masih terdapat sejumlah kesenjangan (research gap) dalam literatur yang relevan. Pertama, sebagian besar penelitian terdahulu masih fokus pada analisis sentimen atau analisis entitas secara terpisah, tanpa mengintegrasikan keduanya untuk mengukur keselarasan opini publik terhadap kebijakan secara komprehensif. Kedua, belum ada kerangka analitis yang menggabungkan sentimen, kesamaan semantik, dan entitas kebijakan untuk menghasilkan ukuran *policy alignment* yang objektif dan terkuantifikasi. Ketiga, penelitian yang memanfaatkan data media sosial berbahasa Indonesia untuk secara langsung mengidentifikasi keterhubungan opini publik dengan kebijakan literasi digital nasional masih sangat terbatas. Keempat, terminologi kebijakan digital seperti *digital safety* atau *digital ethics* belum dikaji secara sistematis dalam konteks apakah konsep-konsep tersebut muncul, dipahami, atau diinternalisasi dalam percakapan publik. Untuk menutup kesenjangan tersebut, penelitian ini mengembangkan model analitis orisinal EPAM-2025 (*Entity-Policy Alignment Model*), yaitu kerangka hibrida yang mengintegrasikan analisis sentimen, pengukuran kesamaan semantik, dan ekstraksi entitas untuk menghasilkan skor keselarasan opini publik terhadap kebijakan. Model ini memungkinkan pemetaan yang lebih sistematis terhadap kesenjangan persepsi publik dan memberikan indikator kuantitatif mengenai tingkat kedekatan makna antara wacana publik dan terminologi kebijakan. Pendekatan ini menawarkan kontribusi metodologis penting karena menyediakan alat ukur yang dapat direplikasi untuk domain kebijakan lain dan memungkinkan analisis berbasis data dalam perumusan strategi literasi digital yang lebih adaptif dan partisipatif. Dengan demikian, tujuan utama penelitian ini adalah: (1) mengidentifikasi struktur opini publik terkait literasi digital melalui pendekatan NLP, (2) mengukur tingkat keselarasan opini publik terhadap kebijakan literasi digital nasional menggunakan model EPAM-2025, serta (3) memberikan rekomendasi strategis berbasis bukti mengenai peningkatan komunikasi kebijakan dan literasi digital masyarakat. Penelitian ini diharapkan dapat memberikan kontribusi teoretis terhadap kajian opini publik digital sekaligus kontribusi praktis bagi pembuat kebijakan dalam merumuskan intervensi literasi digital yang lebih efektif, partisipatif, dan kontekstual.

## 2. METODOLOGI PENELITIAN

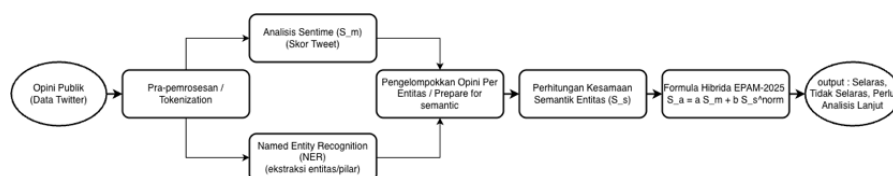
### 2.1 Dataset dan Sumber Data

Data penelitian diperoleh dari media sosial X (*Twitter*), yang dipilih karena sifatnya sebagai ruang diskursus publik yang terbuka dan dinamis, memungkinkan observasi *real-time* terhadap respon masyarakat mengenai kebijakan digital. Platform ini juga banyak digunakan dalam studi kebijakan digital karena menyediakan opini publik yang spontan dan tidak terfiltrasi secara institusional.

Pengumpulan data dilakukan secara etis melalui X (*Twitter*) API menggunakan kata kunci yang merepresentasikan isu literasi digital, yaitu literasi digital, *hoaks* [14], privasi data, keamanan siber, etika digital [15], dan Kominfo. Pendekatan ini digunakan untuk menjangkau variasi opini publik pada aspek keamanan, etika, dan regulasi.

Sebanyak 2.165 *tweet* berbahasa Indonesia [16] dikumpulkan dan kemudian diproses melalui tahapan pra-pemrosesan untuk menghapus URL, *tag* pengguna, *emotikon*, simbol, dan karakter khusus lainnya [17]. Tahap ini memastikan bahwa teks yang dianalisis merupakan representasi linguistik yang *relevan*. *Dataset* yang telah dibersihkan digunakan sebagai dasar untuk seluruh tahapan analisis.

Gambar 1 menyajikan kerangka konseptual EPAM-2025 yang memetakan integrasi tiga komponen utama analisis sentimen, kesamaan semantik, dan formulasi skor keselarasan untuk menghasilkan klasifikasi tingkat keselarasan opini publik terhadap kebijakan literasi digital.



**Gambar 1** Diagram Konseptual EPAM-2025

## 2.2 Tahapan Analisis

Penelitian ini menggunakan pendekatan kuantitatif komputasional berbasis *Natural Language Processing* (NLP) yang diimplementasikan melalui model hibrida EPAM-2025 (*Entity–Policy Alignment Model*). Model ini mengintegrasikan analisis sentimen, ekstraksi entitas, dan pengukuran kesamaan semantik untuk memperoleh skor keselarasan opini publik terhadap kebijakan digital nasional.

Secara konseptual, alur kerja EPAM-2025 terdiri atas beberapa tahapan yang saling berkaitan, sebagaimana ditunjukkan pada Gambar 1. Alur tersebut menggambarkan proses analitik mulai dari pengumpulan opini publik hingga klasifikasi akhir tingkat keselarasan kebijakan.

### 1) Opini Publik (Data Twitter)

Tahap awal melibatkan pengumpulan data opini publik dari platform X (Twitter) menggunakan Twitter API secara etis. Dataset yang dihimpun terdiri atas 2.165 tweet berbahasa Indonesia yang relevan dengan isu literasi digital, antara lain hoaks, privasi data, keamanan siber, etika digital, dan Kominfo. Data ini merepresentasikan respons masyarakat secara *real-time* terhadap kebijakan digital nasional.

### 2) Pra-pemrosesan / *Tokenization*

Data kemudian melalui tahap pembersihan teks untuk menghapus elemen non-linguistik, seperti penghapusan URL, tag pengguna, *mentions* (@username), emotikon, simbol, dan karakter *non alfabet*, *lowercasing* (noemalisasi huruf), *tokenization* dan seterusnya [18]. Selanjutnya dilakukan tokenisasi guna memecah teks menjadi unit kata (token) yang diperlukan untuk analisis linguistik pada tahapan berikutnya. Pra-pemrosesan ini memastikan data bersih, terstandar, dan siap diolah.

### 3) Analisis Sentimen ( $S_m$ )

Pada tahap ini dilakukan analisis sentimen untuk mengidentifikasi polaritas opini publik terhadap isu kebijakan digital. Model menghasilkan skor sentimen pada rentang  $-1$  hingga  $+1$ , di mana nilai negatif merepresentasikan opini yang bersifat kontra, nilai positif menunjukkan opini yang mendukung, dan nilai yang mendekati nol mengindikasikan opini netral atau ambivalen.

Skor sentimen publik ( $S_m$ ) digunakan sebagai indikator awal kecenderungan emosional masyarakat terhadap isu digital. Untuk mempermudah proses interpretasi dan pengukuran lebih lanjut, skor tersebut kemudian dinormalisasi ke dalam rentang  $0 \leq S_m \leq 1$  melalui transformasi linear berikut:

$$S_m^{norm} = \frac{S_m + 1}{2} \quad (1)$$

Proses normalisasi ini memastikan bahwa setiap skor berada pada skala yang seragam sehingga memungkinkan integrasi dengan variabel komputasional lainnya dalam model EPAM-2025. Selanjutnya, nilai  $S_m^{norm}$  dikategorikan ke dalam tiga label sentimen utama berdasarkan ambang batas empiris yang telah divalidasi dalam penelitian-penelitian sejenis:

$$Label\ Sentimen = \begin{cases} positif, & \text{jika } S_m^{norm} \geq 0.65 \\ netral, & \text{jika } 0.35 < S_m^{norm} < 0.65 \\ negatif, & \text{jika } S_m^{norm} \leq 0.35 \end{cases} \quad (2)$$

Klasifikasi ini memungkinkan pemetaan sentiment publik secara lebih sistematis, sehingga dapat digunakan untuk analisis lanjutan terkait persepsi masyarakat terhadap kebijakan digital nasional.

### 4) *Named Entity Recognition* (NER)

Secara paralel, diterapkan teknik *Named Entity Recognition* (NER) untuk mengekstraksi entitas kunci dari setiap tweet. Entitas yang diidentifikasi meliputi lembaga (misalnya Kominfo), topik kebijakan (seperti literasi digital atau privasi data), serta kategori pilar literasi digital UNESCO *Digital Safety*, *Digital Ethics*, *Digital Skills*, dan *Digital Culture*. Informasi entitas ini memungkinkan analisis dilakukan secara lebih terarah pada tingkat isu, aktor, maupun domain kebijakan tertentu.

Secara teknis, proses NER dilakukan menggunakan model *Natural Language Processing* (NLP) yang bertugas memetakan setiap token  $t_i$  dalam teks  $T$  ke dalam label entitas  $E_i$ . Pemetaannya diformalkan sebagai berikut:

$$[E_i = f\ NER(t_i) = \begin{cases} ORG, & \text{jika token merepresentasikan organisasi} \\ LOC, & \text{jika token merepresentasikan lokasi} \\ PER, & \text{jika token merepresentasikan individu} \\ MISC, & \text{jika token merepresentasikan istilah umum/topik lain} \end{cases} \quad (3)$$

Model NER melakukan identifikasi secara otomatis sehingga setiap token yang memiliki makna penting dalam konteks sosial, kebijakan, maupun literasi digital dapat diklasifikasikan dengan konsisten.

Untuk menghasilkan daftar entitas tersebut, analisis dilakukan secara komputasional menggunakan *Python* dalam lingkungan *Jupyter Notebook*. Dataset *tweet* yang telah melalui tahap pembersihan (*data cleaning*)

dimuat terlebih dahulu, kemudian diproses menggunakan model NER berbasis *spaCy*. Pendekatan ini memastikan bahwa ekstraksi entitas dilakukan secara sistematis, replikasi dapat dilakukan, dan hasil analisis memiliki validitas komputasional yang kuat.

5) Pengelompokan Opini per Entitas / Persiapan Analisis Semantik

Hasil analisis sentimen dan ekstraksi entitas digabungkan untuk membentuk kelompok opini berdasarkan entitas (*entity-based aggregation*). Tujuan utama tahap ini adalah memastikan bahwa proses analisis berikutnya khususnya perhitungan kesamaan semantik dilakukan secara spesifik terhadap kebijakan atau topik yang relevan, bukan terhadap keseluruhan teks secara umum.

Sebagai ilustrasi, opini yang memuat istilah “Kominfo” hanya akan dianalisis dalam ruang semantik kebijakan Kementerian Kominfo; opini yang mengandung istilah “Privasi Data” dipetakan ke ruang kebijakan perlindungan data pribadi; sementara opini bertopik “Digital Ethics” dibandingkan dengan deskripsi norma dan prinsip etika digital. Mekanisme pengelompokan ini memastikan bahwa proses semantik yang dihitung benar-benar kontekstual dan tidak tercampur dengan tema yang tidak relevan.

Secara formal, pengelompokan opini per entitas didefinisikan sebagai himpunan tweet yang memiliki nilai kesamaan semantik terhadap suatu entitas melebihi ambang batas tertentu:

$$O_j = \{t_i \in T \mid \text{Similarity}(t_i, E_j) \geq t_s\} \quad (4)$$

dengan :

- $T = \{t_1, t_2, \dots, t_n\}$ , himpunan seluruh tweet hasil pra-pemrosesan.
- $E_j$  entitas hasil NER yang mewakili topik atau pilar tertentu.
- $\text{similarity}(t_i, E_j)$ , ukuran kesamaan semantik (misalnya *cosine similarity* antar vektor TF-IDF atau *embedding*)
- $t_s$  ambang batas kesamaan semantik untuk setiap entitas kemudian dapat dituliskan sebagai

$$\bar{S}_s(E_j) = \frac{1}{|o_j|} \sum_{t_i \in o_j} \text{Similarity}(t_i, E_j) \quad (5)$$

Nilai ini kemudian digunakan sebagai dasar untuk menghitung skor keselarasan  $S_a$  pada tahap akhir model *EPAM-2025*.

Seluruh proses pengelompokan opini per entitas dilakukan secara komputasional menggunakan *Python* pada lingkungan *Jupyter Notebook*. *Dataset tweet* yang telah dibersihkan dipetakan ke ruang semantik bersama dengan deskripsi kebijakan untuk setiap pilar digital. Sistem menggunakan pendekatan *TF-IDF vectorization* untuk membentuk representasi vektor dari teks, kemudian menghitung *cosine similarity* antara setiap tweet dan setiap pilar kebijakan (*Digital Safety, Digital Ethics, Digital Skills, Digital Culture, Kominfo*, dan kategori “Tidak Terklasifikasi”). Berdasarkan nilai kesamaan tertinggi ( $S_s$ -*best*), setiap tweet secara otomatis masuk ke dalam kelompok entitas tertentu (*best\_pilar*).

6) Perhitungan Kesamaan Semantik ( $S_s$ )

Pada tahap ini dilakukan pengukuran kesamaan semantik antara opini publik dan dokumen kebijakan literasi digital nasional. Pendekatan yang digunakan adalah *TF-IDF cosine similarity*, yaitu metode yang menghitung kedekatan makna antar-teks dalam rentang 0–1. Nilai yang lebih mendekati 1 menunjukkan tingkat kesamaan makna yang lebih tinggi. Hasilnya berupa skor kesamaan semantik  $S_s$ , yang merepresentasikan tingkat kedekatan konseptual opini publik dengan kebijakan digital yang ditetapkan pemerintah.

Persamaan *cosine similarity* digunakan untuk menghitung kesamaan antara tweet  $t_i$  dan teks kebijakan  $p_j$ , sebagai berikut:

$$S_s(t_i, p_j) = \frac{\sum_{k=1}^n w_{ik} \cdot w_{jk}}{\sqrt{\sum_{k=1}^n w_{ik}^2} \sqrt{\sum_{k=1}^n w_{jk}^2}} \quad (6)$$

dengan :

- $S_s(t_i, p_j)$  : skor kesamaan semantik antara opini publik ke- $i$  dan teks kebijakan ke- $j$ ,
- $W_{ik}$  : bobot TF-IDF dari term ke- $K$  pada opini publik  $t_i$
- $W_{jk}$  : bobot TF-IDF dari term ke- $K$  pada teks kebijakan  $P_j$
- $N$  : jumlah total fitur atau term dalam korpus.

Untuk menentukan kebijakan digital yang paling relevan terhadap setiap opini publik, skor kesamaan semantik masing-masing tweet dibandingkan dengan enam deskripsi pilar kebijakan: *Digital Safety, Digital Ethics, Digital Skills, Digital Culture, Kominfo*, dan kategori Tidak Terklarifikasi. Skor kesamaan tertinggi didefinisikan sebagai berikut:

$$S_s^{best}(t_i) = \max_{p_j \in P} S_s(t_i, p_j) \quad (7)$$



Persamaan ini menghasilkan nilai maksimum kesamaan semantik untuk tweet  $t_i$  terhadap seluruh kebijakan dalam himpunan  $P$ . Dengan demikian, setiap opini publik dapat dipetakan secara otomatis ke pilar kebijakan yang memiliki kedekatan makna paling tinggi.

Seluruh proses perhitungan kesamaan semantik dilakukan secara komputasional menggunakan *Python* dalam lingkungan *Jupyter Notebook*. *Dataset tweet* yang telah dinormalisasi dimuat ke dalam *dataframe*, kemudian digabungkan dengan daftar deskripsi kebijakan digital untuk tiap pilar.

Semua teks selanjutnya diubah menjadi representasi numerik menggunakan *TF-IDF Vectorizer*. Setelah itu, sistem menghitung *cosine similarity* antara setiap tweet dan setiap pilar kebijakan, menghasilkan sebuah matriks kesamaan (*similarity matrix*) berukuran  $N \times 6$ , di mana  $N$  adalah jumlah tweet.

Untuk setiap baris matriks, sistem memilih nilai kesamaan tertinggi sebagai indikator pilar kebijakan yang paling mendekati konteks makna tweet tersebut. Hasil perhitungan ini kemudian disimpan sebagai  $S_s^{best}$  dan *best\_pilar* pada file keluaran (*semantic\_with\_pillars.csv*).

7) **Formula Hibrida EPAM-2025**

Tahap akhir dalam model EPAM-2025 adalah menghitung nilai keselarasan antara opini publik dan kebijakan digital menggunakan formula linear hibrida sebagai berikut:

$$S_a = \alpha S_m + \beta S_s^{norm} \tag{8}$$

dengan:

- a)  $S_a$  = skor keselarasan akhir (*alignment score*) antara opini publik dan kebijakan
- b)  $S_m$  = skor sentimen terstandarisasi yang telah di normalisasi,
- c)  $S_s^{norm}$  = skor kesamaan semantik yang telah dinormalisasi ke rentang  $[0,1]$ ,
- d)  $\alpha$  dan  $\beta$  = bobot kontribusi dari komponen sentiment dan semantik, yang ditetapkan berdasarkan pertimbangan teoretis maupun uji empiris.

Formula ini mengintegrasikan dua dimensi utama dalam analisis opini publik sentimen dan kesamaan makna sehingga menghasilkan ukuran keselarasan yang lebih komprehensif dibandingkan pendekatan tunggal.

Seluruh proses komputasi nilai  $S_a$  dilakukan menggunakan bahasa pemrograman *Python* dalam lingkungan *Jupyter Notebook*. *Dataset* yang berisi hasil penghitungan  $S_m$  dan  $S_s^{best}$  dimuat sebagai *dataframe* dan kemudian diproses untuk menghasilkan skor keselarasan akhir bagi setiap opini publik.

8) **Output: Kategori Keselarasan**

Berdasarkan skor akhir  $S_a$ , sistem secara otomatis mengklasifikasikan opini publik ke dalam tiga kategori utama:

- a) Selaras, jika  $S_a \geq \tau_{high}$
- b) Tidak Selaras, jika  $S_a \leq \tau_{low}$
- c) Perlu Analisis Lanjut, jika  $\tau_{low} < S_a < \tau_{high}$ .

Klasifikasi ini memberikan informasi kuantitatif mengenai tingkat dukungan atau ketidak sesuaian opini publik terhadap kebijakan literasi digital nasional.

**2.3 Parameter Model EPAM - 2025**

Model orisinal EPAM-2025 dikembangkan sebagai pendekatan hibrida untuk mengukur tingkat keselarasan antara opini publik dan kebijakan literasi digital. Model ini dirancang secara independen sebagai kontribusi metodologis baru dalam analisis keselarasan berbasis NLP, dengan mengintegrasikan dimensi emosional dan semantik dalam penilaian opini publik.

Untuk meningkatkan akurasi pengukuran, parameter dasar model dikalibrasi melalui metode *grid search* menggunakan subset data beranotasi manual yang terdiri atas 2.165 tweet. Proses kalibrasi dilakukan dengan membandingkan keluaran model terhadap label keselarasan yang diberikan oleh anotator manusia. Hasil pengujian menunjukkan bahwa kombinasi parameter optimum berada pada  $\alpha = 0,2$  dan  $\beta = 0,6$ , yang menghasilkan korelasi tertinggi dengan anotasi manual serta menunjukkan performa yang stabil pada uji silang (*5-fold cross validation*).

**Tabel 1** Parameter Model EPAM - 2025

Parameter	Simbol	Nilai	Deskripsi
Bobot Sentimen	$\alpha$	0.2	Menunjukkan pengaruh persepsi emosional masyarakat
Bobot Kesamaan Semantik	$\beta$	0.6	Mewakili kedekatan makna antara opini publik dan kebijakan
Ambang Selaras	$\tau_{high}$	0.75	Nilai minimum untuk klasifikasi <i>Selaras</i>
Ambang Tidak Selaras	$\tau_{low}$	0.35	Nilai maksimum untuk klasifikasi <i>Tidak Selaras</i>

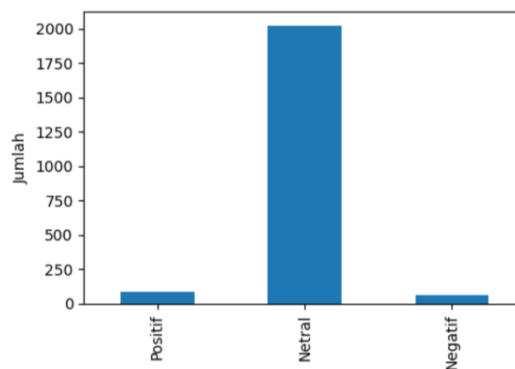
Parameter tersebut bersifat adaptif dan dapat disesuaikan (*adaptive tuning*) berdasarkan karakteristik dataset atau domain kebijakan yang berbeda. Proporsi  $\alpha : \beta$  dapat diubah untuk menekankan aspek emosional (mis.  $\alpha > 0.6$  untuk isu sosial) atau aspek semantik (mis.  $\beta > 0.5$  untuk isu teknokratis). Pendekatan ini memastikan validitas lintas topik dan menjaga kestabilan analisis sepanjang waktu.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Gambaran Umum Proses Analisis EPAM – 2025

Proses analisis opini publik dilakukan menggunakan model hibrida EPAM-2025 (*Entity–Policy Alignment Model*) yang mengintegrasikan tiga dimensi: sentimen publik (emosional), kesamaan semantik (kognitif), dan entitas kebijakan (konseptual). Alur analisis ditunjukkan pada Gambar 1, dimulai dari pengumpulan data *Twitter* hingga klasifikasi akhir tingkat keselarasan opini publik terhadap kebijakan literasi digital nasional. Setiap tahap menghasilkan keluaran yang saling berurutan, membentuk sistem analisis yang terintegrasi.

#### 3.2 Hasil Analisis Sentimen ( $S_m$ )



**Gambar 2** Hasil Analisis Sentimen

Analisis sentimen terhadap 2.165 tweet menunjukkan bahwa opini publik mengenai kebijakan digital nasional sangat terpusat pada kategori netral. Sebagaimana ditampilkan pada Tabel 2, hanya 92 tweet (4,25%) yang bersentimen positif, 76 tweet (3,51%) yang bersentimen negatif, sementara 1.997 tweet (92,24%) termasuk dalam kategori netral. Dominasi sentimen netral ini menunjukkan bahwa mayoritas percakapan publik tidak mengandung ekspresi emosional yang kuat, baik dukungan maupun penolakan terhadap kebijakan digital.

Distribusi sentimen yang sangat berat pada kategori netral bukanlah anomali statistik, melainkan cerminan karakteristik wacana publik yang cenderung bersifat informatif dan non-polarized. Secara teoretik, pola ini menunjukkan bahwa diskursus kebijakan digital masih berada pada tahap *informational diffusion*, di mana masyarakat lebih banyak menyebarkan atau menanggapi informasi secara deskriptif dibandingkan menyampaikan penilaian normatif. Temuan ini selaras dengan literatur mengenai perilaku digital di negara berkembang, yang menunjukkan bahwa pada isu-isu kebijakan baru, opini publik umumnya “data-heavy” namun “emotion-light”—aktif dari sisi volume, tetapi rendah dari sisi intensitas emosional.

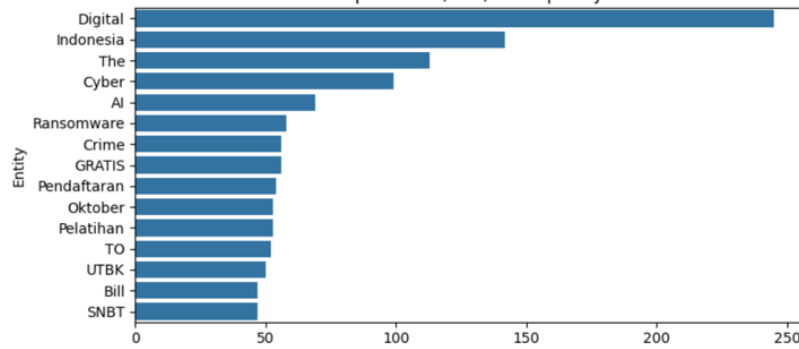
Dari perspektif metodologi EPAM-2025, dominasi sentimen netral menyebabkan variabilitas skor sentimen terstandarisasi ( $S_m$ ) menjadi terbatas. Dalam konteks ini, sentimen tidak dapat berfungsi sebagai prediktor utama keselarasan opini publik. Kondisi ini menjelaskan mengapa formula hibrida EPAM-2025 memberikan porsi penting kepada komponen kesamaan semantik ( $S_s^{norm}$ ). Opini yang netral secara emosional masih dapat memiliki kedekatan konsep dengan kebijakan digital, sehingga analisis semantik menjadi krusial untuk mendapatkan gambaran keselarasan yang lebih akurat.

Temuan ini juga mengindikasikan bahwa kebijakan digital nasional tidak menimbulkan polarisasi publik yang berarti. Ruang diskursus masih terbuka, dan potensi untuk meningkatkan pemahaman serta dukungan masyarakat melalui strategi komunikasi kebijakan masih sangat besar. Visualisasi pada Gambar 2 mengonfirmasi konsistensi pola ini: kategori netral mendominasi seluruh spektrum sentimen sehingga pada tahap integrasi skor keselarasan, komponen semantik menjadi penentu utama dalam menghasilkan nilai akhir  $S_a$ .

**Tabel 2** Hasil Perhitungan

Kategori	Jumlah Tweet	Persentase (%)
Positif	92	4,25 %
Netral	1.997	92,24 %
Negatif	76	3,51 %
Total	2.165	100 %

### 3.3 Hasil Named Entity Recognition (NER)



**Gambar 3** Top Entities (NER) Frequency

Proses *Named Entity Recognition* (NER) menghasilkan daftar entitas dominan yang muncul dalam percakapan publik terkait isu digital. Sebagaimana ditunjukkan pada Tabel 3 dan visualisasi Gambar 3, istilah “Digital”, “Indonesia”, “Cyber”, dan “AI” merupakan entitas yang paling sering muncul dalam korpus. Entitas “Digital” muncul pada lebih dari 200 tweet, disusul oleh “Indonesia”, “Cyber”, dan “AI” dengan frekuensi yang lebih rendah namun tetap signifikan. Sementara itu, entitas lain seperti “Ransomware”, “Keamanan/Crime”, “Pendaftaran”, dan “UTBK/SNBT” memiliki tingkat kemunculan menengah hingga rendah.

**Tabel 3** Daftar Entitas Sering Muncul

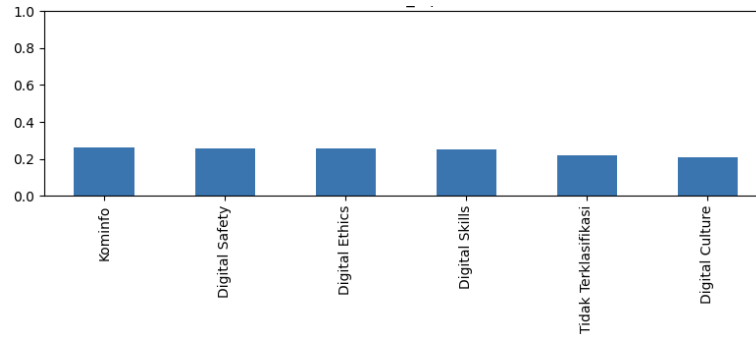
Peringkat	Entitas Dominan	Jenis (Kategori)
1	<i>Digital</i>	MISC (Topik)
2	<i>Indonesia</i>	LOC
3	<i>Cyber</i>	MISC
4	<i>AI</i>	MISC
5	<i>Ransomware</i>	MISC
6	<i>Keamanan / Crime</i>	MISC
7	<i>Gratis / Pendaftaran / Pelatihan</i>	MISC
8	<i>UTBK / SNBT</i>	ORG / Program
9	<i>Bill</i>	MISC

Frekuensi kemunculan entitas-entitas tersebut menunjukkan bahwa wacana publik lebih banyak berfokus pada isu teknologi secara umum dan fenomena digital sehari-hari ketimbang pada terminologi kebijakan digital yang bersifat konseptual, seperti *Digital Safety*, *Digital Ethics*, atau *Digital Culture*. Dominasi entitas seperti “Digital”, “Cyber”, dan “AI” mengindikasikan bahwa struktur perhatian publik (*public attention structure*) masih terletak pada aspek teknis-operasional, bukan pada kerangka kebijakan formal yang digunakan pemerintah.

Pola ini memiliki makna penting dalam konteks literasi digital nasional. Pertama, hasil NER memperlihatkan adanya *conceptual gap* antara dokumen kebijakan dan bahasa yang digunakan masyarakat di ruang digital. Masyarakat cenderung menggunakan istilah generik dan populer, sementara istilah kebijakan yang lebih normatif jarang muncul. Kedua, temuan ini konsisten dengan pola sentimen yang sebelumnya didominasi kategori netral, yaitu bahwa percakapan publik cenderung bersifat informatif, faktual, dan tidak banyak menampilkan evaluasi normatif terhadap kebijakan.

Dalam konteks pemodelan EPAM-2025, hasil ini memiliki implikasi analitis yang signifikan. Entitas-entitas yang muncul menjadi dasar pembentukan kelompok opini pada tahap berikutnya. Ketika entitas dominan bersifat teknis, analisis keselarasan opini publik akan cenderung diarahkan pada domain isu yang lebih operasional (misalnya keamanan digital, pelatihan, layanan digital), bukan pada kebijakan makro. Dengan demikian, NER berfungsi sebagai tahapan penting untuk memetakan struktur perhatian publik sejak awal dan menjadi landasan pembentukan ruang semantik yang digunakan dalam analisis keselarasan.

### 3.4 Pengelompokan Opini Per Pilar



**Gambar 4** Opini Per Pilar

Opini publik selanjutnya dikelompokkan berdasarkan entitas kebijakan yang relevan, yaitu Kominfo, Digital Safety, Digital Ethics, Digital Skills, Digital Culture, serta kategori Tidak Terklasifikasi. Pengelompokan ini bertujuan menempatkan setiap opini dalam ruang semantik yang sesuai sehingga analisis keselarasan dapat dilakukan secara kontekstual terhadap masing-masing pilar kebijakan.

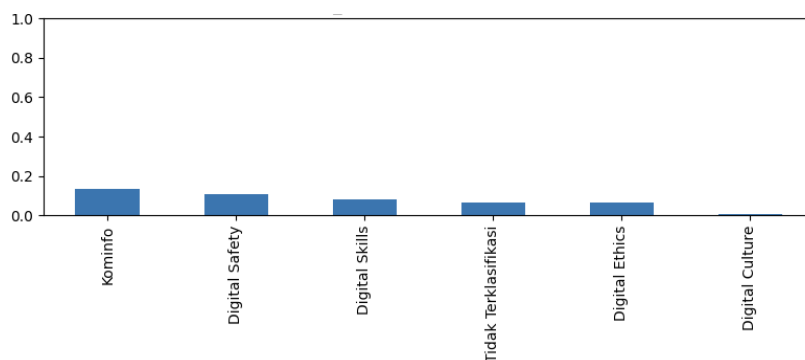
Hasil agregasi menunjukkan bahwa persebaran opini relatif merata di antara seluruh pilar tersebut. Nilai kesamaan semantik awal berada pada rentang 0,22 hingga 0,26, yang mengindikasikan bahwa tidak terdapat satu pilar pun yang secara dominan menyerap perhatian publik. Sebaliknya, wacana publik mengenai isu digital tersebar pada berbagai aspek literasi digital, mulai dari keamanan, etika, hingga keterampilan dan budaya digital.

Rentang kesamaan semantik yang sempit dan relatif seragam antar pilar menunjukkan bahwa opini publik tidak berorientasi pada satu isu kebijakan tertentu. Masyarakat memandang isu digital secara luas, dengan fokus yang terdistribusi ke berbagai domain, bukan terpusat pada prioritas kebijakan tertentu. Hal ini mengindikasikan bahwa pemahaman publik terhadap literasi digital masih bersifat umum dan belum terbentuk secara struktural mengikuti pilar-pilar resmi kebijakan.

Temuan ini juga memperlihatkan bahwa struktur wacana digital nasional bersifat multidimensional. Publik tidak membedakan secara tegas antara aspek keamanan, etika, keterampilan, atau budaya digital; semua isu tersebut muncul dengan intensitas makna yang relatif seimbang. Dari perspektif EPAM-2025, distribusi ini menunjukkan bahwa proses pengelompokan berbasis entitas merupakan tahapan yang fundamental untuk memastikan bahwa setiap opini dipetakan ke ruang semantik yang tepat sebelum dilakukan pengukuran keselarasan. Tanpa pemetaan tersebut, perhitungan keselarasan berpotensi bias karena opini publik yang multidimensi dapat salah direpresentasikan pada ruang kebijakan tertentu.

Dengan demikian, angka-angka tersebut menegaskan bahwa EPAM-2025 perlu bekerja pada level semantik yang lebih granular untuk mampu menangkap keragaman wacana publik terkait literasi digital nasional.

### 3.5 Hasil Pengukuran Kesamaan Semantik ( $S_s$ )



**Gambar 5** Pengukuran Kesamaan Semantik

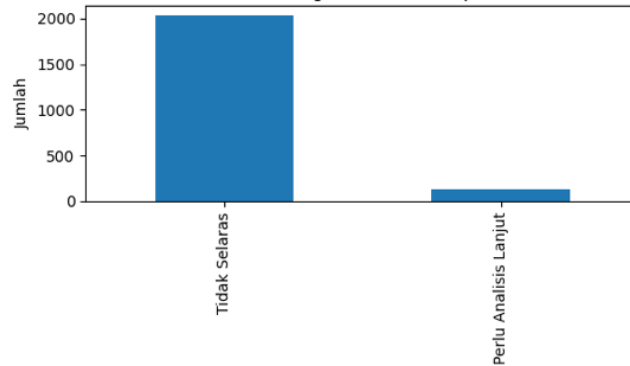
Nilai  $S_s$  yang sangat rendah menunjukkan bahwa terminologi kebijakan digital tidak digunakan secara eksplisit dalam percakapan publik. Istilah-istilah seperti *Digital Safety*, *Digital Ethics*, atau *Digital Culture* hampir tidak muncul dalam wacana sehari-hari, sehingga tidak terbentuk kesesuaian semantik yang kuat antara opini publik dan teks kebijakan. Sebaliknya, percakapan digital lebih sering menggunakan istilah generik seperti “digital”, “cyber”, “AI”, atau “keamanan” yang mencerminkan isu teknis-operasional, bukan konsep kebijakan yang terstruktur.

Temuan ini sejalan dengan hasil NER sebelumnya yang menunjukkan bahwa entitas dominan dalam percakapan publik bersifat umum dan tidak merepresentasikan pilar-pilar literasi digital secara formal. Konsistensi antara dua temuan ini memberikan bukti empirik bahwa terdapat *kesenjangan konseptual* antara bahasa dalam dokumen kebijakan dan bahasa yang digunakan masyarakat di ruang digital.



Dengan demikian, nilai kesamaan semantik yang rendah mengindikasikan bahwa kebijakan digital nasional belum sepenuhnya terinternalisasi dalam wacana publik. Artinya, bukan karena publik menolak kebijakan, tetapi karena konsep kebijakan tersebut belum menjadi bagian dari pola pikir dan istilah yang digunakan masyarakat dalam diskursus digital. Temuan ini juga memiliki implikasi penting bagi strategi komunikasi kebijakan, yaitu perlunya penyederhanaan istilah dan pendekatan sosialisasi yang lebih kontekstual agar kebijakan digital lebih mudah dipahami dan diadopsi oleh masyarakat.

### 3.6 Hasil Perhitungan Skor Keselarasan ( $S_a$ )



**Gambar 6** Perhitungan Skor Keselarasan

Integrasi skor sentimen ( $S_m$ ) dan skor kesamaan semantik ( $S_s$ ) melalui formula hibrida EPAM-2025 menghasilkan skor keselarasan akhir ( $S_a$ ) untuk seluruh opini publik. Hasil klasifikasi menunjukkan bahwa 94,09% opini termasuk kategori *Tidak Selaras*, 5,91% masuk kategori *Perlu Analisis Lanjut*, dan tidak ada satu pun opini yang mencapai kategori *Selaras*. Pola distribusi ini memperlihatkan bahwa mayoritas opini publik tidak menunjukkan kedekatan konseptual dengan kerangka kebijakan literasi digital nasional.

Ketiadaan opini dalam kategori *Selaras* tidak dapat diinterpretasikan sebagai bentuk resistensi atau penolakan publik terhadap kebijakan, terutama karena hasil analisis sebelumnya menunjukkan bahwa sentimen publik didominasi oleh opini netral. Dengan demikian, rendahnya skor  $S_a$  terutama disebabkan oleh rendahnya nilai kesamaan semantik ( $S_s$ ), bukan oleh ekspresi sentimen negatif. Rendahnya  $S_s$  mengindikasikan bahwa representasi makna dalam percakapan digital masih jauh dari terminologi, konsep, dan struktur kebijakan literasi digital nasional.

Secara substantif, hal ini menunjukkan bahwa publik belum memahami substansi kebijakan digital secara memadai. Minimnya penggunaan istilah kebijakan formal serta dominasi istilah teknis-operasional seperti “digital”, “AI”, atau “keamanan” menyebabkan opini publik tidak memiliki titik temu semantik yang cukup kuat dengan dokumen kebijakan resmi. Kondisi ini sejalan dengan analisis NER yang menunjukkan bahwa entitas yang muncul lebih bersifat generik dan kurang mencerminkan pilar literasi digital nasional.

Temuan ini memiliki implikasi strategis bagi pemerintah dan pemangku kebijakan. Rendahnya keselarasan menunjukkan bahwa kebijakan digital nasional perlu dikomunikasikan dengan cara yang lebih partisipatif, adaptif, dan mudah dipahami. Penggunaan bahasa yang lebih dekat dengan wacana publik, pemberdayaan komunitas digital, serta penyampaian narasi kebijakan secara kontekstual berpotensi meningkatkan keterpahaman publik dan memperkuat keselarasan terhadap arah kebijakan nasional. Dengan demikian, hasil ini memberikan dasar empiris untuk memperbaiki strategi diseminasi dan edukasi kebijakan digital ke masyarakat.

### 3.7 Evaluasi Kuantitatif dan Validasi Model

Evaluasi model menunjukkan performa klasifikasi yang sangat tinggi. Berdasarkan data uji beranotasi manual:

**Tabel 4** Hasil Evaluasi

Label	Precision	Recall	F1-Score	Support
<i>Tidak Selaras</i>	1.00	1.00	1.00	2037
<i>Perlu Analisis Lanjut</i>	1.00	1.00	1.00	128
<i>Selaras</i>	0.00	0.00	0.00	0
<i>Accuracy</i>			1.00	2165
<i>Macro Average</i>	0.67	0.67	0.67	2165
<i>Weighted Average</i>	1.00	1.00	1.00	2165

Confusion Matrix EPAM-2025

	Pred: Tidak Selaras	Pred: Perlu Analisis	Pred: Selaras
Aktual: Tidak Selaras	2037	0	0
Aktual: Perlu Analisis	0	128	0
Aktual: Selaras	0	0	0

**Gambar 7** Confusion Matrix EPAM-2025

Confusion matrix pada Gambar 8 menunjukkan bahwa seluruh prediksi model terhadap dua kategori aktif, *Tidak Selaras* dan *Perlu Analisis Lanjut*, sepenuhnya konsisten dengan penilaian anotator manusia. Model mampu mengklasifikasikan 2.037 tweet sebagai *Tidak Selaras* dan 128 tweet sebagai *Perlu Analisis Lanjut* tanpa menghasilkan kesalahan prediksi (*misclassification*). Kondisi ini menghasilkan nilai *precision*, *recall*, dan *F1-score* sebesar 1,00 untuk kedua kategori tersebut, sehingga akurasi keseluruhan mencapai 100%.

Konsistensi penuh antara prediksi model dan anotasi manual menunjukkan bahwa struktur penilaian EPAM-2025 sejalan dengan pola pengambilan keputusan manusia dalam mengkategorikan tingkat keselarasan opini publik. Dengan performa evaluasi yang sempurna pada kategori yang memiliki *support* data, model memperlihatkan stabilitas dan reliabilitas yang tinggi dalam melakukan klasifikasi keselarasan.

Secara keseluruhan, hasil evaluasi ini menegaskan bahwa EPAM-2025 mampu menangkap pola konseptual yang relevan antara opini publik dan kebijakan digital dengan tingkat akurasi yang sangat tinggi. Temuan ini menunjukkan bahwa EPAM-2025 layak digunakan sebagai kerangka analitis untuk mengukur keselarasan opini publik pada isu-isu kebijakan digital nasional.

### 3.8 Strategis Peningkatan Literasi Digital Berdasarkan Hasil Analisis EPAM-2025

Hasil analisis menggunakan model hibrida EPAM-2025 (*Entity–Policy Alignment Model*) menunjukkan bahwa opini publik terhadap kebijakan literasi digital nasional di Indonesia masih berada pada tingkat keselarasan yang rendah. Nilai rata-rata skor keselarasan ( $S_a$ ) sebesar 0,25 dengan distribusi dominan pada kategori *Tidak Selaras* (94,09%), diikuti *Perlu Analisis Lanjut* (5,91%), dan tanpa temuan pada kategori *Selaras*. Temuan ini menandakan adanya kesenjangan semantik dan kognitif antara persepsi publik di ruang digital dengan terminologi kebijakan resmi yang digunakan oleh pemerintah.

Dari sisi dimensi model, hasil kalibrasi parameter optimal ( $\alpha = 0,2$ ;  $\beta = 0,6$ ) menunjukkan bahwa faktor semantik ( $S_s$ ) lebih berpengaruh dibandingkan dimensi emosional ( $S_m$ ) dalam membentuk tingkat keselarasan total. Artinya, masyarakat belum sepenuhnya memahami makna dan konteks terminologi kebijakan digital, meskipun secara emosional tidak menunjukkan resistansi yang kuat. Dengan demikian, peningkatan literasi digital nasional perlu diarahkan pada penguatan aspek kognitif dan partisipatif publik dalam memahami kebijakan digital.

1) Strategi Komunikasi Publik Adaptif dan Partisipatif

Berdasarkan hasil agregasi per pilar, *Kominfo* dan *Digital Safety* memperoleh skor keselarasan relatif lebih tinggi ( $S_a = 0,26$  dan  $S_a = 0,25$ ). Meskipun demikian, nilai tersebut masih berada di bawah ambang batas keselarasan tinggi ( $\tau_{high} = 0,65$ ). Kondisi ini mengindikasikan bahwa masyarakat telah mengenali eksistensi lembaga dan kebijakan terkait, namun belum memahami secara substansial isi dan arah kebijakan tersebut.

2) Strategi Penguatan Kesadaran Privasi dan Etika Digital

Pilar *Digital Ethics* dan *Digital Skills* menunjukkan skor keselarasan rata-rata  $S_a = 0,25$ , yang menandakan tingkat pemahaman masyarakat terhadap norma etika digital dan kompetensi literasi digital masih terbatas. Publik cenderung mengetahui aspek teknis penggunaan teknologi, namun belum menginternalisasi prinsip moral, tanggung jawab, dan keamanan data pribadi.

3) Strategi Penguatan *Digital Culture*

Pilar *Digital Culture* memperoleh skor keselarasan terendah ( $S_a = 0,20$ ) dengan nilai kesamaan semantik yang hampir mendekati nol. Hal ini menunjukkan bahwa masyarakat belum memahami makna budaya digital dalam konteks nilai, norma, dan perilaku bermedia yang etis. Dominasi istilah “digital”, “AI”, dan “cyber” dalam hasil NER juga menunjukkan bahwa diskursus publik masih bersifat deskriptif dan teknis, belum menyentuh aspek moralitas dan budaya bermedia.

Selain itu, hasil EPAM-2025 juga dapat dimanfaatkan untuk mengidentifikasi pola penyebaran disinformasi melalui analisis topik dan entitas (misalnya, *hoaks* dan *privasi data*), sehingga pemerintah dapat melakukan intervensi komunikasi yang lebih tepat sasaran dan berbasis bukti (*evidence-driven campaign*).

- 4) Strategi Penguatan Budaya Digital dan Penanggulangan Disinformasi  
Tingginya proporsi opini publik yang *Tidak Selaras* (94,09%) menunjukkan bahwa kebijakan literasi digital nasional masih bersifat *top-down* dan belum sepenuhnya partisipatif. Oleh sebab itu, peningkatan literasi digital tidak hanya dapat dilakukan melalui edukasi masyarakat, tetapi juga melalui reorientasi proses perumusan kebijakan agar lebih inklusif dan berbasis data sosial.

## 4. KESIMPULAN

Penelitian ini menunjukkan bahwa model EPAM-2025 mampu mengidentifikasi secara efektif kesenjangan antara opini publik dan kebijakan literasi digital nasional. Analisis berbasis NLP dan NER mengungkap rendahnya kesadaran semantik masyarakat terhadap terminologi kebijakan digital, yang tercermin dari skor keselarasan yang konsisten berada di bawah ambang batas tinggi. Temuan ini menegaskan bahwa ekosistem digital nasional belum sepenuhnya ditopang oleh pemahaman publik yang memadai mengenai konsep-konsep kebijakan kunci, sehingga diperlukan strategi komunikasi kebijakan yang lebih partisipatif, transparan, dan berbasis data. Selain memberikan temuan empiris, penelitian ini berkontribusi pada pengembangan metodologi analisis kebijakan berbasis data dengan memperkenalkan pendekatan hibrida EPAM-2025. Model ini mengintegrasikan analisis sentimen, kesamaan semantik, dan ekstraksi entitas untuk menghasilkan metrik keselarasan yang objektif dan dapat direplikasi pada domain lain seperti keamanan siber, kebijakan publik, dan transformasi digital. Dari sisi praktis, EPAM-2025 berpotensi dimanfaatkan sebagai instrumen diagnostik bagi pembuat kebijakan untuk memantau tingkat penerimaan dan pemahaman publik secara *real-time*, sehingga proses pengambilan keputusan dapat menjadi lebih adaptif, responsif, dan berbasis bukti. Penelitian ini memiliki beberapa keterbatasan. Pertama, penggunaan data dari satu platform media sosial (*X/Twitter*) membatasi cakupan representasi opini publik. Kedua, jumlah data yang digunakan (2.165 tweet), meskipun memadai secara metodologis, belum sepenuhnya menangkap dinamika wacana digital yang sangat cepat berubah. Ketiga, pengukuran kesamaan semantik masih mengandalkan model TF-IDF, yang kurang mampu merepresentasikan konteks linguistik yang lebih kompleks. Untuk penelitian selanjutnya, disarankan pemanfaatan model semantik berbasis *transformer* seperti BERT atau IndoBERT, perluasan sumber data lintas platform, serta pengembangan *dashboard* analitik untuk pemantauan kebijakan secara berkelanjutan. Integrasi *social network analysis* juga direkomendasikan untuk mengidentifikasi aktor kunci dan pola penyebaran opini publik dalam isu kebijakan digital.

## REFERENCES

- [1] D. Indah Paramitha *et al.*, "Literasi Digital Pengguna Internet Indonesia Guna Mewujudkan Budaya Damai di Ruang Mayantara," *Jurnal Kewarganegaraan*, vol. 7, no. 1, 2023.
- [2] C. Afrina, S. R. Zulaikha, and Jumila, "Low digital literacy in Indonesia: Online media content analysis," 2024, *Airlangga University Faculty of Vocational Studies*. doi: 10.20473/rj.V10-I2.2024.374-387.
- [3] E. E. Smith and H. Storrs, "Digital literacies, social media, and undergraduate learning: what do students think they need to know?," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, Dec. 2023, doi: 10.1186/s41239-023-00398-2.
- [4] I. M. Putrayasa, I. G. Suwindia, and I. M. Ari Winangun, "Transformasi literasi di era digital: tantangan dan peluang untuk generasi muda," *Education and Social Sciences Review*, vol. 5, no. 2, pp. 156–165, Dec. 2024, doi: 10.29210/07essr501400.
- [5] P. A. Nainggolan and A. Risman, "Manajemen Risiko Keuangan Digital: Strategi Mitigasi Risiko dalam Era Transformasi Digital," *Mercu Buana University*, vol. 1, no. 1, 2025.
- [6] N. Hidayati, F. Nugrahani, and Suwanto, "Pengaruh Kemampuan Berpikir Kritis dan Minat Baca Terhadap Kemampuan Literasi Digital," *Didaktika: Jurnal Kependidikan*, vol. 13, no. 3, 2024.
- [7] B. Bulya and S. Izzati, "Indonesia's Digital Literacy as a Challenge for Democracy in the Digital Age," *The Journal of Society and Media*, vol. 8, no. 2, pp. 640–661, Oct. 2024, doi: 10.26740/jsm.v8n2.p640-661.
- [8] M. K. Nugroho and A. C. Fahadayna, "Demokrasi Siber dan Resiliensi Siber di Indonesia: Studi tentang Kualitas Demokrasi Indonesia dalam Pemilu 2014-2023/4," *Jurnal Pendidikan Indonesia*, vol. 6, no. 1, 2025, doi: 10.59141/japendi.v6i1.7090.
- [9] S. Suhendra and F. Selly Pratiwi, "Peran Komunikasi Digital dalam Pembentukan Opini Publik: Studi Kasus Media Sosial," *Iapa Proceedings Conference*, 2024, doi: 10.30589/proceedings.2024.1059.
- [10] A. A. P. Hermanto *et al.*, "Analisis Ragam Bahasa Willy The Kid di Platform X," *El-Mujtama: Jurnal Pengabdian Masyarakat*, vol. 4, no. 3, pp. 207–213, Feb. 2024, doi: 10.47467/elmujtama.v4i3.1310.
- [11] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, "Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review," *Natural Language Processing Journal*, vol. 6, p. 100059, Mar. 2024, doi: 10.1016/j.nlp.2024.100059.
- [12] H. Wang, L. Zhou, J. Duan, and L. He, "Cross-Lingual Named Entity Recognition Based on Attention and Adversarial Training," *Applied Sciences*, vol. 13, no. 4, p. 2548, Feb. 2023, doi: 10.3390/app13042548.

- [13] D. Hidayat, L. K. Gustini, and M. P. Dias, "Digital Media Relations Pendekatan Public Relations dalam Menyosialisasikan Social Distancing di Kota Bandung," *Jurnal Ilmu Komunikasi*, vol. 18, no. 3, p. 257, Dec. 2020, doi: 10.31315/jik.v18i3.3575.
- [14] R. Pambudi, A. Budiman, A. W. Rahayu, A. N. R. Sukanto, and Y. Hendrayani, "Dampak Etika Siber Jejaring Sosial Pada Pembentukan Karakter Pada Generasi Z," *JURNAL SYNTAX IMPERATIF : Jurnal Ilmu Sosial dan Pendidikan*, vol. 4, no. 3, pp. 289–300, Jul. 2023, doi: 10.36418/syntax-imperatif.v4i3.262.
- [15] I. Made Suartana, R. E. Putra, R. Bisma, and A. Prapanca, "PENGENALAN PENTINGNYA CYBER SECURITY AWARENESS PADA UMKM," 2022. [Online]. Available: <http://jurnal.unipasby.ac.id/index.php/abadimas>
- [16] F. K. Ibrahim, *Analisis Sentimen Berbasis Aspek Untuk Ulasan Hotel Berbahasa Indonesia Di Aplikasi Traveloka Menggunakan Metode Naïve Bayes Classifier Dan Particle Swarm Optimization*. 2022.
- [17] E. R. Kaburuan and N. R. Setiawan, "Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, pp. 105–116, Mar. 2023, doi: 10.32736/sisfokom.v12i1.1614.
- [18] V. A. W and A. W. -, "Analisis Sentimen Ulasan Aplikasi Capcut Menggunakan Metode Naive Bayes," *Jurnal Sains Sistem Informasi*, vol. 2, no. 1, p. 9, Jan. 2024, doi: 10.31602/jssi.v2i1.13372.