INFLUMATCH AI: AI KOL – Product Matching Platform
An AI-Assisted Engine for Influencer Selection

Natchaya Temyignyong[*,1], Wilasinee Paewboontra[*,1], Theeranop Kerdnana[*,2],
Krittanon Kaewtawee[1], Wachiravit Modecrua[1], Thanaphat Suwatanapornchai[2], Apicha Suksompong[2],
Patrapol Pichitphan[2], Touchapon Kraisingkorn[1]

Abstract

Influencer marketing increasingly depends on selecting Key Opinion Leaders (KOLs) whose content, audience, and style align with a product brief, yet manual shortlisting is time-consuming and inconsistent. This paper presents INFLUMATCH AI, an AI-assisted KOL–product matching platform that converts a brand brief into a ranked shortlist with transparent scores and explanations. INFLUMATCH AI adopts a three-agent pipeline: (1) a Product Extracting Agent that normalizes product/campaign details and infers target-audience and KOL preference signals; (2) a KOL Profile Extracting Agent that builds structured KOL profiles by retrieving platform metadata, transcripts, engagement/activity statistics, and content category/style distributions, with optional news/sentiment signals; and (3) a Product $\times$ KOL Matching Agent that embeds preference signals for vector retrieval, generates a tier-aware campaign plan, and applies reasoning-based multi-criteria scoring on retrieved candidates to produce final rankings. Evaluation covers 200 KOLs for profile extraction and 29 matching test cases derived from 14 products. The KOL Profile Extracting Agent achieves 83% pass on demographics and profile/statistics extraction and 72% pass on content analysis. The Product Extracting Agent achieves 84.6% pass. End-to-end matching achieves 80.6% recall (coverage) against human reference shortlists. Results indicate that structured, agentic decomposition can reduce manual effort while maintaining strong alignment with human selections; remaining errors are primarily driven by transcript-only content understanding and incomplete coverage of campaign-specific criteria in the reasoning rubric.

Keywords: Influencer Marketing; KOL Matching; Multi-Agent Systems; Vector Similarity Search; LLM Reasoning

1 Introduction

Over the past few years, consumer media habits have shifted rapidly in the digital era. Traditional cable and satellite television—once a primary channel for brand advertising—has continued to decline as audiences spend more of their attention on social platforms and digital content. Deloitte's *Digital Media Trends* reports highlight this transition: based on online surveys of U.S. consumers, the share of respondents with a cable or satellite TV subscription fell from 63% in 2022 to 49% in 2025 [1, 2]. In parallel, global social media adoption has reached massive scale, with 5.24 billion social media user identities—about 63.9% of the world's population as of February 2025 [3, 8].

This shift is especially pronounced among younger generations. Daily internet usage tends to increase as age decreases, reaching roughly seven hours per day among people aged 16–34 [4, 6, 8]. Gen Z, in particular, lives a mobile-first lifestyle where shopping discovery, evaluation, and purchase decisions increasingly begin online. Recent marketing statistics suggest that 97% of Gen Z consumers use social media as their primary

---
[*]Co-first authors
[1]Amity AI Research and Application Center, {natchaya, wilasinee.pae}@amitysolutions.com
[2]EGG Digital Company Limited, {theeranop.ker}@ascendcorp.com

source of shopping inspiration, while 67.4% rely on online reviews and 66% avoid products with outdated reviews [4]. As social feeds increasingly shape purchase intent, influencer content also plays a stronger role in preference formation—Shopify reports that 36% of respondents state influencer content outperforms brand-created content [5].

Together, these trends indicate a clear shift from brand-only messaging toward KOL-led influence at scale. Unlike traditional brand advertising—which is polished, one-directional, and tightly controlled—KOL marketing operates through ongoing audience relationships [18, 19, 20, 21]. KOLs communicate in a more conversational, platform-native style, embedding product mentions into routines, tutorials, reviews, and live interactions, which often feels more like a trusted recommendation than an interruption [9, 10]. This shift is reinforced by the way social platforms reward authenticity, frequency, and engagement, and by the fact that discovery and evaluation increasingly happen inside feeds, where consumers can see products demonstrated, compared, and discussed across multiple creators [4]. Brands still define positioning and guardrails, but KOLs translate those messages into narratives that resonate with specific micro-communities; the advantage comes from fit—the right KOL, reaching the right audience, in the right style, at the right time.

However, the same factors that make KOL-led influence effective also make it operationally difficult. The KOL ecosystem is large and dynamic, and "fit" depends on far more than follower counts: audience composition, content themes, tone of voice, brand safety, past collaborations, and performance stability all matter. As a result, influencer selection becomes a repeated decision problem that is difficult to execute manually at scale.

To address this, this paper proposes INFLUMATCH AI, an AI-assisted KOL–product matching system that (i) structures unstructured briefs into explicit requirements, (ii) extracts standardized KOL profiles from platform signals, and (iii) matches creators to briefs using a hybrid approach that combines embedding-based semantic retrieval (to efficiently narrow the candidate pool) with LLM-based multi-criteria reasoning and weighted scoring (to rank candidates and generate explanations). This design reduces manual browsing effort, improves consistency, and provides clearer justification for why particular KOLs are recommended.

2 System Architecture & Design

INFLUMATCH AI is an AI-assisted KOL–product matching platform that streamlines influencer selection into a repeatable, data-driven pipeline. Given a brand product/campaign brief, the system outputs a ranked shortlist of KOLs with transparent fit scores and explainable rationales. The design goal is to reduce manual browsing and subjective decision-making by standardizing how campaign requirements and KOL signals are represented, compared, and evaluated.

The system architecture was informed by practitioner research with KOL selection specialists. Interviews and working sessions were conducted to understand (i) how specialists translate a brief into selection requirements, (ii) how they screen and shortlist KOLs, and (iii) which criteria drive final decisions. Across these discussions, influencer selection consistently emerged as a multi-stage workflow: specialists first clarify the brief and define constraints, then explore and filter candidate creators, and finally perform deeper evaluation based on content relevance, audience fit, tone/style alignment, past collaborations, and performance stability. This practitioner workflow was analyzed and abstracted into a pipeline that preserves the most decision-critical steps while reducing redundant effort through automation.

2.1 Inputs

INFLUMATCH AI operates on three inputs:

- A Product/Campaign Brief B

- An optional KOL Preference Brief P that specifies desired KOL characteristics

- A KOL Candidate Set $K = \{k_1, \ldots, k_n\}$ that provides a pool of KOLs to consider

2.1.1 Handling Missing Inputs (Auto-Analysis)

Because real-world briefs are often incomplete, INFLUMATCH AI supports partially specified inputs: Brand/Product Name is the only mandatory field. When optional fields are omitted, the system infers missing attributes from the available product text and attachments (if any) and proceeds with extraction and matching. When P is not provided, the system defaults to ranking based on brief-derived intent, emphasizing relevance, audience-fit proxies, and safety/risk signals.

2.2 Outputs

The platform returns a ranked list of top-m recommended KOLs R. For each candidate in R, the system provides:

- Fit score breakdown across key dimensions (e.g., relevance, audience fit, tone alignment, performance proxies, safety/risk)

- Constraint status (pass/fail for hard requirements)

- Explanations (reason codes and supporting evidence snippets used in the decision)

These outputs support rapid shortlisting and stakeholder justification (e.g., "why this KOL was recommended").

2.3 Structured Representations

To enable consistent matching, INFLUMATCH AI converts raw inputs into structured schemas:

- **Brief schema** S_B**:** a normalized representation of the product/campaign requirements. It separates hard constraints (must be satisfied) from soft preferences (optimized during ranking), and encodes explicit priorities/weights when provided or inferred.

- **KOL schema** S_{k_i}**:** a normalized representation of each KOL. It aggregates extracted KOL information into a unified profile that may include identity and metadata, content/category and style descriptors, audience proxies (when available), time-windowed activity and performance statistics, and optional risk-related signals.

This structured layer reduces ambiguity from free-form inputs and enables direct comparison across candidates.

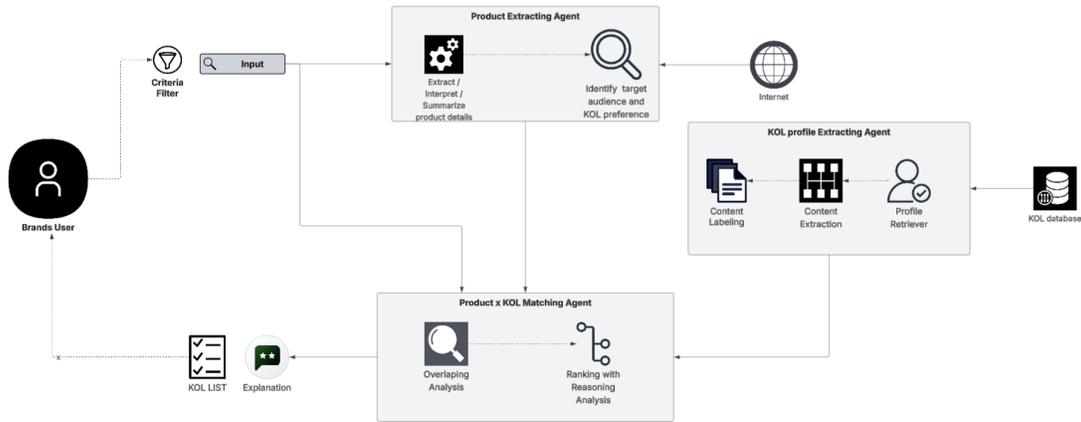## 2.4 Execution Flow and Components

Figure 1: System overview of INFLUMATCH AI architecture.

As illustrated in Figure 1, the platform is implemented as a modular three-stage pipeline:

1. **Product Extracting Agent:** transforms the Product/Campaign brief B (and optional preference brief P) into the structured brief schema S_B. When necessary, it enriches product understanding using available contextual resources (e.g., attached documents and additional product descriptions).

2. **KOL Profile Extracting Agent:** transforms each candidate k_i into a structured KOL schema S_{k_i} by extracting KOL metadata and observable signals from the underlying profile/source, including bio information, location/language cues, recent content, content themes, and engagement/activity statistics (e.g., followers, likes, comments, posting frequency). Where available, it also captures collaboration-related indicators.

3. **Matching Agent (Product × KOL):** performs (i) constraint filtering, (ii) scoring and ranking, and (iii) explanation generation to produce the final shortlist R.

The pipeline is intentionally modular: extraction and decisioning are separated to improve maintainability, debugging, and interpretability.

2.5 Product Extracting Agent

The Product Extracting Agent converts a brand's product/campaign brief into a structured representation that can be used downstream for matching. In addition to parsing information explicitly provided in the brief, the agent retrieves supplementary public context (e.g., product positioning, category cues, and common use cases) to improve completeness when the brief is sparse. The agent outputs two artifacts: (i) structured product details $S_{B_{\text{prod}}}$ and (ii) inferred target audience $S_{B_{\text{aud}}}$ and KOL preference signals $S_{B_{\text{pref}}}$ that guide the matching stage.
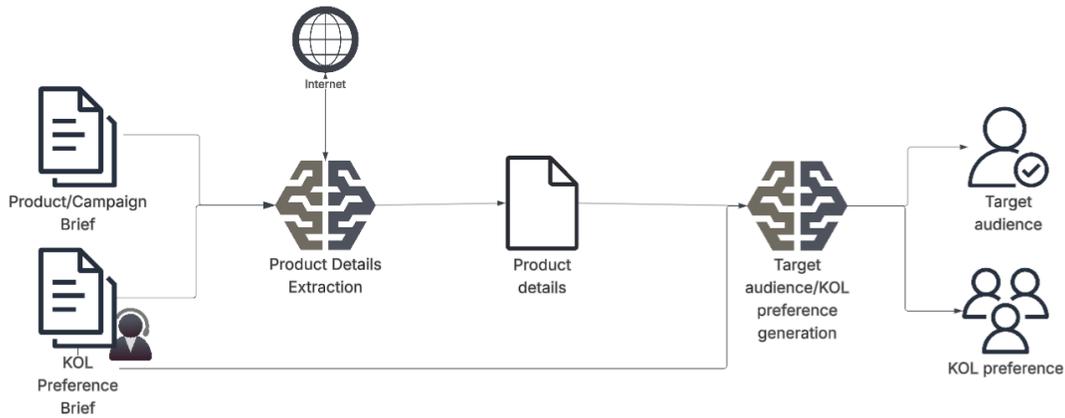
Figure 2: Product Extracting Agent workflow.

As illustrated in Figure 2, the Product Extracting Agent consists of two steps.

2.5.1 Step 1: Product Details Extraction

This step extracts and normalizes product information from the input brief. When the brief lacks sufficient detail, the system optionally performs web retrieval to enrich the description with publicly available information relevant to the product. In the implementation, this stage uses GPT-5 mini[11, 12] with a web search tool for retrieval-augmented extraction.

Input

- Product/Campaign brief B

Output

- Product Details Schema $S_{B_{\mathrm{prod}}}$

The resulting structured product details reduce ambiguity and provide consistent signals for downstream audience and KOL preference inference.

2.5.2 Step 2: Target Audience and KOL Preference Generation

Given the extracted product details, the agent infers the most likely target audience for the product and derives an initial KOL preference profile suitable for promoting and representing the product. If the user provides a KOL Preference Brief P, it is incorporated as constraints or priors; otherwise, the agent generates preferences purely from product context. In the implementation, this stage uses GPT-5 mini[11, 12].

Inputs

- Product Details Schema $S_{B_{\text{prod}}}$

- KOL Preference Brief P (optional)

Outputs

- Target Audience Profile $S_{B_{\text{aud}}}$

- KOL Preference Profile $S_{B_{\text{pref}}}$

These outputs serve as the brief-side representation for later stages. They inform constraint filtering when applicable and provide soft preference signals used by the Matching Agent to score KOL–product fit.

2.6 KOL Profile Extracting Agent

The KOL Profile Extracting Agent transforms each candidate KOL $k_i \in K$ (initially provided as a handle, URL, or platform identifier) into a structured profile S_{k_i}. The agent is responsible for retrieving observable KOL signals, deriving time-windowed performance features, and summarizing content characteristics (category and style distributions) in a standardized format for downstream matching.

In the implementation, the agent integrates the TikWM API for TikTok profile and video retrieval[17],[16] the Supadata API for transcript extraction, web search for recent news retrieval, and Anthropic Claude for vision-based demographic inference and text-based content classification[14].

Input

- k_i: KOL identifier (e.g., TikTok handle, channel name, or user ID)

Output
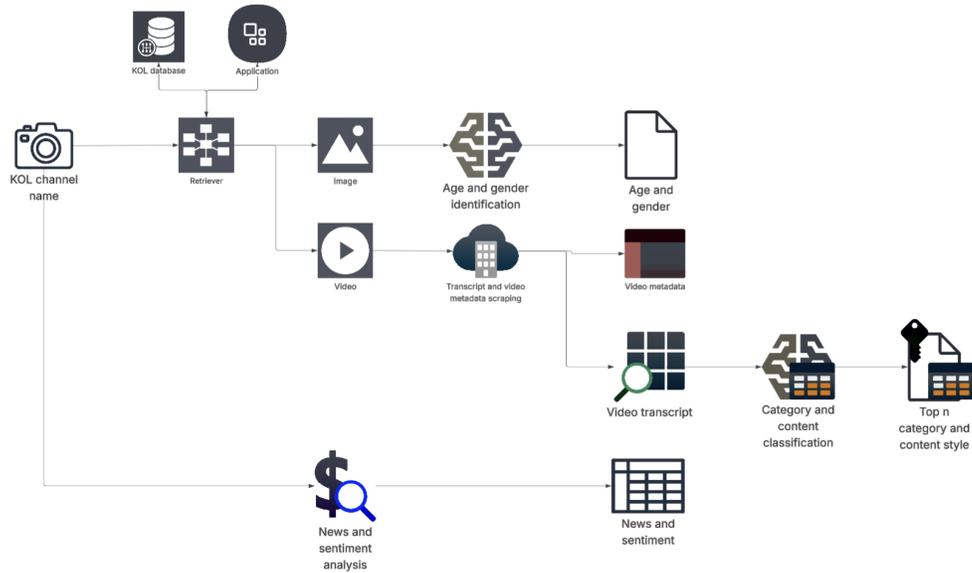
- S_{k_i}: structured KOL schema

Figure 3: KOL Profile Extracting Agent workflow.

As illustrated in Figure 3, the KOL Profile Extracting Agent consists of five steps.

2.6.1 Step 1: Profile Photo Analysis

The agent first checks the platform's internal database for self-declared demographic attributes (e.g., gender and age) for KOLs who have registered and provided profile information. If these attributes are missing, unavailable, or not publicly accessible, the agent falls back to downloading the KOL's profile photo from TikTok and inferring coarse demographic cues. Profile images are retrieved using TikWM API (or direct URL retrieval when available)[17]. Demographic inference is performed using Anthropic Claude Vision[15], which outputs gender labels and age ranges.

2.6.2 Step 2: Video and Transcript Retrieval

For each KOL, the agent retrieves recent TikTok videos within a fixed window (e.g., the last 30 days) using TikWM API[17]. It extracts per-video metadata and engagement signals—likes, shares, saves, comments, views, duration, and publish time—and derives temporal features such as day-of-week and posting time buckets. Video transcripts are obtained via the Supadata API[16]. When certain engagement statistics are incomplete (e.g., comment counts), the agent optionally enhances them using lightweight web scraping built on requests.

2.6.3 Step 3: AI-Based Content Classification (Category and Style)

The agent applies transcript-based classification to characterize each KOL's content. For each video transcript, Anthropic Claude is used to predict (i) a category, and (ii) a content style. Each classifier returns the

7

top-3 labels with associated probabilities. The agent then aggregates predictions across videos to compute KOL-level distributions (e.g., top categories/styles with percentages).

2.6.4 Step 4: News and Sentiment Analysis (Risk Signal)

To capture potential reputational risk signals, the agent optionally performs a recent-news scan (e.g., last 30 days) for each KOL using web search. Retrieved mentions are summarized and assigned a sentiment label (positive/neutral/negative) using Anthropic Claude[15]. This step produces a compact representation of notable events, controversies, or public criticism when such signals are detectable.

2.6.5 Step 5: Profile Statistics Aggregation and Feature Construction

Finally, the agent aggregates profile- and video-level signals into standardized features. It computes time-windowed engagement metrics (7/14/30 days), including engagement rate and average likes/shares/saves/comments/views, as well as activity indicators such as posting frequency, churn duration (days since last post), and average posting interval. It also derives posting pattern features such as the most common time slot and day of week.

2.7 Product $\times$ KOL Matching Agent

The Product $\times$ KOL Matching Agent is responsible for converting brief-side intent into a ranked, explainable shortlist of KOLs. It integrates three key capabilities: (i) semantic retrieval over KOL embeddings to efficiently narrow the candidate space, (ii) strategic tier planning to determine how many KOLs to select per tier under campaign constraints, and (iii) LLM-based reasoning and scoring to produce final rankings with transparent justifications.

Inputs

- Product Details Schema $S_{B_{\text{prod}}}$

- Target Audience Profile $S_{B_{\text{aud}}}$

- KOL Preference Profile $S_{B_{\text{pref}}}$

- KOL Preference Brief P (optional)

- KOL candidate pool K

Outputs Tier-separated ranked results with:

- Selected KOLs (top results)

- Alternative KOLs (non-selected candidates)

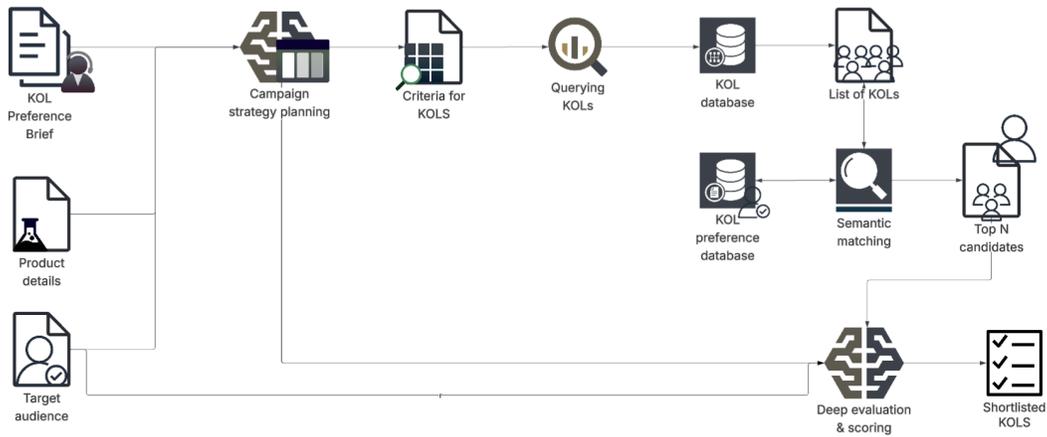- Per-KOL score (0–100) and reasoning summary

Figure 4: Product × KOL Matching Agent workflow.

As illustrated in Figure 4, the Product × KOL Matching Agent consists of four steps.

2.7.1 Step 1: KOL Preference Embedding

To enable semantic retrieval, the Matching Agent converts preference signals into vector embeddings. Preference features are aggregated into a canonical text/feature bundle. The aggregated preference bundle is embedded using the Gemini Embedding API, producing a numpy embedding vector used for similarity search.

2.7.2 Step 2: Campaign Strategy Planning

Given product context, audience profile, and user constraints (e.g., age/gender/tier/budget), the agent generates a tier-level campaign plan that specifies:

- Tiers (e.g., Mega, Macro, Micro, Nano)

- Number of KOLs per tier

- Budget allocation per tier

- Inferred KOL demographic preferences

- Rationale for the proposed plan

This planning stage ensures the final shortlist is consistent with campaign constraints and budget structure.

2.7.3 Step 3: Overlapping and Ranking with Reasoning Analysis

The Matching Agent executes tier-level matching in parallel with each tier following the same two-stage process:

Step 3.1: Semantic Search (Tier-Filtered Retrieval)   Using the preference embedding vector and the tier constraints from the planning output, the system performs vector similarity search:

- Filter candidates by tier (Mega, Macro, Micro, Nano)

- Compute cosine similarity between preference vectors and stored KOL profile embeddings

- Retrieve top-K candidates per tier (default $K = 50$)

- Retrieve associated metadata required for scoring (e.g., profile attributes and content-category matches)

Step 3.2: AI Reasoning and Structured Evaluation   After semantic retrieval returns the top-K candidates per tier, the system applies GPT-5 [11, 12] to perform fine-grained assessment. This stage functions as a quality filter: it validates semantic matches using multi-criteria judgment and produces a structured, auditable evaluation for each candidate.

To control latency and context length, candidates are processed in fixed-size batches (e.g., 10 KOLs per batch) and evaluated in parallel.

For each KOL, the reasoning module returns a structured evaluation record containing:

- KOL name

- Binary selection decision

- Criterion-level evaluations (e.g., a list of eight per-criterion assessments with scores and short justifications)

- Overall justification explaining why the KOL is selected or not selected

- Conflict-of-interest notes highlighting potential concerns (e.g., competitor mentions or mismatched brand associations)

This design ensures that each recommendation can be traced back to explicit criterion-level judgments rather than a single opaque score.

Step 3.3: Weighted Scoring and Selection   Once all chunks complete, the system computes a final weighted score for each candidate by combining the criterion-level scores produced by the reasoning module:

$$\text{Score}(k_i) = \sum_{j=1}^{J} w_j \, s_j(k_i) \tag{1}$$

where $s_j(k_i)$ is the LLM-produced score for criterion j, and w_j is the corresponding criterion weight.

In addition to producing a numeric score, the reasoning module outputs an explicit select or not-select decision for each candidate. For each tier, the system forms the final shortlist by (i) retaining only candidates marked as selected, (ii) sorting them by weighted score in descending order, and (iii) returning the shortlisted top-N KOLs specified by the campaign plan. Candidates not included in the top-N are retained as alternative options.

2.7.4 Step 4: Tier Aggregation and Logging

After tier-level matching completes (executed in parallel across tiers), the system aggregates all results into a final response by merging evaluations across chunks and tiers, ranking candidates within each tier, and compiling structured explanations and conflict-of-interest notes. In addition, the system aggregates execution statistics such as total token usage across chunk evaluations and runtime summaries (e.g., maximum processing time across tiers) for monitoring and profiling.

3 Experimental Setup

INFLUMATCH AI is evaluated component-wise and end-to-end using a combination of human verification and retrieval-style metrics.

3.1 Evaluation Datasets

KOL Profile Extracting Agent
A sample of 200 KOLs is used to assess the accuracy and completeness of extracted KOL attributes.

Product Extracting Agent + Matching Agent (Product $\times$ KOL)
A total of 39 test cases are constructed from 14 products, where products may appear under different tier and budget configurations (e.g., varying tier mixes or constraints). Each test case is evaluated against a candidate pool of approximately 1,000 KOLs, which is used to measure both brief extraction quality and end-to-end matching performance.

Data availability note
The matching stage is supported by an internal KOL database that contains attributes not consistently available from public sources. This provides broader coverage and more stable feature availability for evaluation, but it also introduces a dependency on internal data availability when comparing against methods that rely strictly on public-only signals.

3.2 Evaluation Metrics

3.2.1 Product Extracting Agent

Each test case is manually reviewed and labeled as correct (1) or wrong (0) based on whether the extracted product information and inferred brief-side signals are acceptable under the evaluation rubric. Let T be the number of test cases and let $y_t \in \{0, 1\}$ be the human label for test case t. Accuracy is computed as:

$$\text{Acc}_{\text{prod}} = \frac{1}{T} \sum_{t=1}^{T} y_t \tag{2}$$

3.2.2 KOL Profile Extracting Agent

Each evaluated KOL profile is manually reviewed and labeled as correct (1) or wrong (0) based on whether the extracted profile meets the verification criteria (e.g., key attributes correctly extracted or inferred). Let N be the number of evaluated KOLs and let $Z_i \in \{0, 1\}$ be the human label for KOL i. Accuracy is computed as:

$$\text{Acc}_{\text{kol}} = \frac{1}{N} \sum_{i=1}^{N} Z_i \tag{3}$$

3.2.3 Product $\times$ KOL Matching Agent

For each test case t, let G_t be the set of reference KOLs selected by humans and let R_t be the set of KOLs recommended by the system (i.e., selected results). The number of correct KOLs is $|R_t \cap G_t|$. Case-level coverage is defined as:

$$\text{Cov}_t = \frac{|R_t \cap G_t|}{|G_t|} \tag{4}$$

The final metric reports the average across all test cases:

$$\text{Cov} = \frac{1}{T} \sum_{t=1}^{T} \text{Cov}_t \tag{5}$$

This corresponds to the average fraction of human-referenced KOLs that are recovered by the system per test case.

3.3 Practical Note on Ground Truth Stability

KOL profiles, content, and engagement signals evolve over time; therefore, a human shortlist annotated at one point may become partially outdated when re-evaluated later. To preserve realism while retaining a consistent reference point, the human shortlist is used as a reference for expected coverage, but evaluation does not rely solely on exact list overlap.

Instead, each system-recommended KOL is additionally assessed for brief-consistency at evaluation time (i.e., whether the KOL satisfies the campaign intent and constraints) rather than only checking whether the system's shortlist matches the historical human-selected list exactly.

4 Results

This section reports evaluation outcomes for the three major components: the KOL Profile Extracting Agent, the Product Extracting Agent, and the end-to-end Product $\times$ KOL Matching Agent, as summarized in Table 1.

Table 1: Performance summary of INFLUMATCH AI agents.

Agent	Result
Product Extracting Agent	Pass: 84.6% Pass with condition: 5%
KOL Profile Extracting Agent (Demographics and profile/statistics extraction)	Pass: 83% Pass with condition: 11%
KOL Profile Extracting Agent (Content analysis and classification)	Pass: 72% Pass with condition: 8%
Product × KOL Matching Agent	Recall (Coverage): 80.6%

4.1 Product Extracting Agent

The Product Extracting Agent is evaluated by human verification on whether extracted product information and inferred brief-side signals are acceptable. Results suggest that the agent generally produces usable structured product details and target audience and KOL preference signals, with a small fraction requiring minor clarification or additional context.

4.2 KOL Profile Extracting Agent

Performance is reported across two evaluation categories:

- **Demographics and profile/statistics extraction** (e.g., demographics, follower-level signals, activity statistics)

- **Content analysis and classification** (e.g., top-3 categories and top-3 content styles with proportions)

Overall, the agent achieves higher reliability on structured metadata and numeric signals than on content-derived classifications, which are inherently more sensitive to transcript quality and ambiguity in short-form content.

4.3 Product × KOL Matching Agent

End-to-end matching performance is evaluated using the average coverage metric (recall-like) against human reference shortlists. The result indicates that, on average, the system retrieves and selects a substantial portion of human-referenced KOLs per test case while operating under dynamic KOL signals.

Note "Pass with condition" indicates outputs that are acceptable for downstream use but require minor review (e.g., partial missing fields, borderline classification confidence, or small inconsistencies that do not invalidate the overall profile or recommendation).

5 Discussion

The evaluation results indicate that the proposed pipeline is effective overall, but several limitations emerged that explain most observed failure cases and suggest clear directions for improvement.

5.1 Limitations in content understanding for KOL profiling

The KOL Profile Extracting Agent currently derives content categories and styles primarily from video transcripts. This design is efficient and scalable, but it introduces errors when transcripts are not representative of the visual content. A common failure mode occurs when videos contain music-only segments, background audio, or minimal spoken narration; in such cases the transcript provides little semantic signal, causing the system to misclassify the KOL's content themes. In addition, the current pipeline does not analyze visual cues (frames/video) and therefore cannot reliably capture appearance-driven requirements (e.g., "female with long hair") or other visual attributes that may be important for certain campaigns. These limitations likely contribute to the lower pass rate observed in content analysis compared to structured metadata extraction.

5.2 Sensitivity to brief completeness and ambiguity

Because most brief fields are optional, the quality of downstream outputs depends strongly on the completeness and specificity of the input. When users provide only a minimal cue such as a brand name, the Product Extracting Agent may infer the wrong product identity—especially for ambiguous names shared by multiple products or categories (e.g., "SCOTCH" as tape vs. health beverage). This ambiguity propagates through the pipeline, affecting inferred target audience, KOL preference signals, and final ranking. These cases highlight the importance of incorporating stronger disambiguation mechanisms (e.g., requesting one additional attribute when ambiguity is detected, or using retrieval to confirm product category before inference).

5.3 Coverage of reasoning criteria in matching

The Matching Agent's reasoning module uses a fixed set of criteria to evaluate shortlisted KOLs. While this provides structure and explainability, the current criteria set does not fully cover all needs that may appear in real briefs. Certain campaign-specific constraints—such as particular visual requirements, niche brand-safety considerations, or highly specific content formats—may not be adequately represented in the evaluation rubric. As a result, the reasoning stage may score candidates favorably even when they miss important brief-specific requirements. Expanding the criteria set (or dynamically selecting criteria based on the extracted brief) would improve robustness and alignment with diverse campaign intents.

Implications and future improvements

These issues suggest two practical improvements. First, incorporating multimodal analysis (image/video signals) into the KOL profiling stage would reduce transcript-only failures and enable appearance-driven matching. Second, moving from a fixed reasoning rubric to a more comprehensive and adaptable set of evaluation criteria would improve alignment with the diversity of constraints that brands may specify. Together, these changes would improve reliability and matching quality, particularly in campaigns where visual cues and nuanced brand requirements are critical.

6 Conclusion

This paper presented INFLUMATCH AI, a multi-agent system for KOL–product matching that converts a brand brief into a ranked shortlist of KOLs with transparent scores and explanations. The proposed architecture decomposes the task into three agents: a Product Extracting Agent that structures product and campaign intent, a KOL Profile Extracting Agent that builds standardized KOL profiles from platform signals, and a Matching Agent that combines embedding-based retrieval with reasoning-based scoring to produce tier-aware recommendations.

Experimental results show that the system produces generally reliable outputs across components: the Product Extracting Agent achieved a 84.6% pass and a 5% pass with condition rate, while the KOL Profile Extracting Agent achieved 83% pass and 11% pass with condition for demographics/profile statistics and 72% pass and 8% pass with condition for content analysis. End-to-end matching achieved 80.6% recall (coverage) against human reference shortlists, indicating that the approach can recover a substantial portion of relevant KOLs while reducing manual effort.

The discussion highlighted key limitations, particularly transcript-only content understanding and incomplete coverage of evaluation criteria in the reasoning stage. Future work should focus on incorporating multimodal signals to better capture visual attributes and strengthening the reasoning criteria to support a wider range of campaign requirements. Overall, the results suggest that a structured, multi-agent pipeline is a practical and effective direction for scalable, explainable influencer selection.

References

[1] China Widener, Jana Arbanas, Doug Van Dyke, Chris Arkenberg, Bree Matheson, and Brooke Auxier, "2025 Digital Media Trends: Social platforms are becoming a dominant force in media and entertainment," *Deloitte Insights*, March 25, 2025.

[2] Kevin Westcott, Jana Arbanas, Chris Arkenberg, Brooke Auxier, Jeff Loucks, and Kevin Downs, "2022 Digital media trends, 16th edition: Toward the metaverse," *Deloitte Insights*, March 28, 2022.

[3] Simon Kemp, "Digital 2025: Global Overview Report," *DataReportal – Global Digital Insights*, February 5, 2025.

[4] Ansherina Opena, "Top Generation Z Marketing Statistics [2025] for Marketers," *Cropink*, updated April 24, 2025.

[5] Elise Dopson, "28 Important Influencer Marketing Statistics To Know in 2025," *Shopify Blog*, November 11, 2024.

[6] Aljay Ambos, "Top Gen Z Marketing Behavior Statistics 2025," *Amra and Elma*, July 16, 2025.

[7] Backlinko Team, "Social Media Usage & Growth Statistics," *Backlinko*, updated Nov. 19, 2025.

[8] Telefónica, "How much time do we spend on our digital lives?," *Telefónica*, October 22, 2025.

[9] Edelman, "2019 Edelman Trust Barometer Special Report: In Brands We Trust?," *Edelman*, June 2019.

[10] GWI, "Exploring the current influencer marketing landscape and best practices," *GWI*, 2019.

[11] OpenAI, "Pricing," *OpenAI API Documentation* (model pricing table including `gpt-5-mini`).

[12] OpenAI, "Compare models (GPT-5)," *OpenAI API Documentation*.

[13] Google, "Embeddings," *Gemini API Documentation (Google AI for Developers)*.

[14] Anthropic, "Vision," *Claude Developer Platform Documentation*.

[15] Anthropic, "Web search tool," *Claude Developer Platform Documentation*.

[16] Supadata, "Web & Video to text API for makers," *Supadata*.

[17] TikWM, "Tiktok Scraper," *RapidAPI* (TikWM-provided TikTok scraping endpoints listing).

[18] Martha Kendall, "29 influencer marketing statistics to guide your brand's strategy in 2025," *Sprout Social Insights*, March 3, 2025.

[19] Digital Marketing Institute, "20 Surprising Influencer Marketing Statistics," *Digital Marketing Institute*, April 16, 2025.

[20] Aspire, "The State of Influencer Marketing 2025," *Aspire*, 2025.

[21] Influencer Marketing Hub, "Influencer Marketing Benchmark Report 2024," *Influencer Marketing Hub*, 2024.