



BAX 423 Final Project Report

Artelligence: AI-Powered Creative Marketplace

**Stephen Louie
Medha Dobhal
Sravya Bhaskara
Chaitanya (CK) Khot
Nivi Balasubramanian**

Date: June 3, 2025

Table of Contents

1. Business Objective	2
2. Key Actionable Business Initiative	2
3. Metrics of Success	3
4. Role of Analytics	4
5. Thinking Through the Analytics	5
6. Executing the Analytics	6
7. Implementation	6
8. Scale	7

1. Business Objective

Artelligence is designed to create an intelligent marketplace connecting brands with content creators in a highly personalized, data-driven way. Brands today struggle to identify creators whose visual style, audience profile, and creative strengths truly align with their campaign objectives. Conversely, creators face difficulty finding consistent, well-compensated partnerships that genuinely match their talents. Current platforms frequently overlook critical matching factors like audience alignment, content style, and campaign effectiveness, limiting potential outcomes for both parties.

Our objective is to resolve these inefficiencies by streamlining the entire brand and creator collaboration workflow. By leveraging advanced analytics and generative AI, Artelligence enables precise matching, automatic campaign ideation, and unified messaging. Ultimately, Artelligence will deliver stronger campaign ROI for brands and stable, rewarding opportunities for creators.

2. Key Actionable Business Initiative

Initiatives:

- Automated Brand and Creator Matching: Implemented AI profiling based on engagement data, content style, and audience alignment to identify optimal partnerships.
- Generative AI for Creative Content: Used large language models (LLMs) to automatically generate campaign briefs, social copy, and advertising ideas, minimizing creative barriers and expediting content development.
- Graph Retrieval-Augmented Generation (GraphRAG): Leveraged graph-based recommendations to uncover indirect yet valuable connections, improving match accuracy beyond traditional similarity scoring.
- User-friendly Web Interface: Developed a Streamlit-based web application for brands to input campaign goals and immediately received matched creators and generated content ideas.

Most Impactful Initiative:

The GraphRAG-powered matching system is our core initiative. By identifying indirect, high-value connections through advanced analytics, GraphRAG significantly enhances partnership relevance and campaign effectiveness.

Execution Plan:

- Integrated influencer data from Instagram with synthetic brand profiles using PySpark.
- Engineered key features: engagement metrics, content topics, follower demographics, and brand attributes.
- Built recommendation models combining direct similarity scoring and GraphRAG analytics.
- Deployed functionality through a streamlined, easy-to-use Streamlit application.

Why GraphRAG Matters

Traditional influencer matching often relies solely on direct metrics like follower count, basic engagement scores, or broad content categories. This surface-level approach frequently leads to predictable and repetitive recommendations, limiting creativity and impact.

Graph Retrieval-Augmented Generation (GraphRAG) offers a fundamentally different method by analyzing complex, multi-dimensional relationship networks among creators, brands, and audiences. Instead of simple attribute matching, GraphRAG considers historical collaborations, indirect audience overlaps, and nuanced stylistic affinities. This advanced reasoning enables Artelligence to identify creators who, despite not appearing as obvious candidates through direct metrics, possess high strategic alignment potential.

Without GraphRAG, Artelligence's matching algorithm would lose substantial analytical depth, resulting in generic, repetitive recommendations. The platform would overlook hidden, yet powerful connections such as a creator whose audience profile closely matches a brand's target demographics or whose past collaborations indirectly align with a brand's campaign objectives. GraphRAG ensures recommendations are contextually informed, strategic, and uniquely valuable, enabling brands and creators to pursue more innovative, impactful partnerships.

3. Metrics of Success

Artelligence's effectiveness will be measured by three essential metrics:

- **Match Quality Score (MQS):** Measures alignment between creators and brand campaign objectives based on direct feedback, campaign outcomes, and alignment modeling.

- Campaign Activation Rate: Tracks the percentage of recommendations converting into actual campaigns, reflecting real-world value.
- Content Turnaround Time: Monitors the efficiency of campaign content delivery from brief generation to final content approval.

Metric Prioritization:

The primary metric is Match Quality Score, followed by Campaign Activation Rate and Content Turnaround Time. These reflect the platform's core value proposition and efficiency gains.

Hypothesis:

Artelligence anticipates at least a 30% improvement in match quality compared to traditional matching methods. This enhancement will likely lead to a 20% higher activation rate and reduce content turnaround by 40–50%.

4. Role of Analytics

Analytics forms the essential backbone of Artelligence's strategic capabilities. First, analytics enables large-scale data integration, bringing together diverse datasets from influencers and brands. Without robust analytics, the process of combining, structuring, and analyzing data from varied sources—such as engagement metrics, audience demographics, and content types—would be impractical at scale.

Second, analytics drives continuous innovation and refinement within the platform. By examining patterns of successful and unsuccessful campaigns, analytics identifies critical insights and opportunities for improvement. For example, analysis of historically high-performing campaigns helps refine the matching logic and informs how generative AI prompts are structured. This ensures recommendations and creative outputs remain relevant, timely, and targeted.

Third, analytics facilitates rigorous performance measurement and iterative improvement. Each critical metric (MQS, activation rate, turnaround time) is systematically tracked, visualized, and analyzed using analytics dashboards. This continuous measurement allows Artelligence to validate its underlying hypotheses, quickly identify areas requiring enhancement, and drive informed decisions to optimize business outcomes.

Furthermore, analytics specifically enables advanced features like GraphRAG. Traditional data processing alone cannot adequately interpret the intricate relationships and nuanced patterns inherent in influencer marketing. Analytics, particularly graph analytics, uniquely supports the complex reasoning required to identify subtle yet highly valuable creator-brand relationships. Thus, analytics not only

supports operational tasks but also underpins the platform's strategic differentiation.

5. Thinking Through the Analytics

The analytics process leverages both existing and synthetic datasets for robust performance.

Data Sources:

- Instagram influencer data including bios, content styles, and engagement metrics.
- Synthetic brand profiles generated by LLMs to provide diverse and comprehensive brand scenarios.

Target Outcomes and Features:

The main objective is to recommend optimal creator matches for specific brand campaigns. Primary features include content niche, engagement levels, style, and brand attributes.

Data Variation:

Acknowledging variation in creator style, engagement patterns, and brand profiles, feature engineering methods standardize and normalize these variables for effective modeling.

Analytics Approaches:

- Predictive Analytics: Drives primary matching decisions.
- Exploratory Analytics: Facilitates creative content generation and trend detection.
- Causal Analytics: While primarily predictive, the implied hypothesis is that well-matched partnerships will enhance campaign outcomes.

Potential Challenges:

- Data quality and completeness.
- Accurate interpretation of nuanced brand and creator styles through LLMs.
- Scalability and computational efficiency, addressed through cloud infrastructure.

6. Executing the Analytics

Artelligence's analytics execution follows a structured approach, emphasizing both technical precision and practical usability.

Data Integration and Feature Engineering:

Initial stages involve robust PySpark pipelines to ingest, preprocess, and clean influencer and brand datasets, extracting structured features (engagement metrics, demographics) and unstructured features (content themes, brand tone).

Matching Models:

Analytics execution incorporates two modeling approaches:

- Basic similarity scoring for initial matches.
- Advanced GraphRAG analytics for uncovering deeper indirect relationships, leveraging historical campaign data and network connections to provide insightful, differentiated recommendations.

Content Generation and Delivery:

LLMs generate creative briefs and campaign ideas, automatically integrated into a streamlined web interface accessible to brands in real time.

4. LLM Content Generation (Google Gemini)

Generate engaging content for your campaigns using Google Gemini.

Generate Campaign Content

Select Brand for Content Generation:

Airbnb

Enter Creator Name:

TravelBuddy

Enter Creator Main Category (e.g., 'Travel', 'Fashion'):

Travel

Maximum Number of Words:

20

Content Type:

Instagram Post

Generate Content

Generated Content:

Dream destinations? Airbnb makes it easy – explore local vibes & unforgettable stays! 🌍🏡

Importance of GraphRAG:

Without GraphRAG, Artelligence would lack depth, relying on simplistic matching and missing strategic partnership opportunities. GraphRAG adds essential context, enabling recommendations that significantly exceed conventional expectations.

Graph Construction & Retrieval:

We represent Instagram creators and brands as nodes in a graph, with edges weighted by relevance. Edge Weights are computed using engagement ratios and category similarity. When a brand queries the system, the graph retrieves top-matching creators by traversing these weighted connections.

These top nodes serve as context for LLM-based content generation.

6. Graph-based Influencer Recommendation (Graph RAG) 🔗

This section uses a graph database concept to find the best influencers or celebrities you can collaborate with for best value, based on engagement metrics and category relevance.

Find Best Influencers for a Brand

Select a Brand:

Airbnb

Number of top influencers to recommend:

15

120

Get Recommendations

Top 5 Influencers for Airbnb:

	Influencer Name	Subscribers	Category 1
0	london	2196385	travel
1	mauronakada	1488478	travel
2	sjanaelise	1480938	travel
3	taramilkea	1168914	travel
4	andreadipietro10	787201	travel

7. Implementation

The implementation of Artelligence involves clear phases focused on technical reliability, analytical rigor, and user-centricity. Initially, technical infrastructure is established using cloud platforms like Google Cloud Platform (GCP), enabling scalable data handling, automated pipelines, and computational power for sophisticated analytics.

Model development and validation occur iteratively, ensuring accurate and reliable recommendations through rigorous testing against historical data and performance benchmarks. User feedback plays a pivotal role at every stage, informing feature adjustments, usability improvements, and overall system reliability.

The Streamlit-based web application is designed through collaborative input from brands and creators, ensuring intuitive navigation and ease of use. Implementation includes comprehensive testing, pilot launches, and careful monitoring of system performance and user satisfaction.

Effective communication channels, user support, and training resources ensure smooth adoption. Implementation success is systematically evaluated, with continuous enhancements based on real-time analytics insights, user feedback, and evolving market requirements.

8. Scale

Artelligence's scalability strategy encompasses both technological and operational dimensions. Technically, the architecture leverages scalable cloud computing infrastructure, ensuring seamless expansion of data storage, processing capabilities, and real-time analytics. PySpark enables distributed data processing, maintaining performance even as data volume significantly increases.

Strategically, scaling involves incorporating additional data sources from emerging platforms such as YouTube, TikTok, and Twitter, broadening the platform's analytical reach and enhancing its competitive advantage. Regularly integrating new features, such as real-time trend analysis and multi-platform content management, maintains market relevance and ensures the platform continues to meet evolving user needs.

Operational scalability is supported through comprehensive user onboarding, accessible educational resources, and robust community support systems. Continuous performance tracking via analytics dashboards informs strategic decisions, ensuring Artelligence remains agile, responsive, and capable of swiftly adapting to market shifts and user feedback.

By carefully balancing technical infrastructure, strategic growth, and proactive user support, Artelligence is well positioned to sustainably scale, maintain its innovative edge, and continually deliver meaningful value to brands and creators in the evolving digital marketplace.