

Facebook Post Scraper & Analyzer (Handover Manual & Walkthrough)

This document provides a comprehensive overview and user guide for the Facebook Post Scraper system. It is designed to assist marketing agencies in deriving social media strategies by automating the scraping, analysis, and organization of Facebook content.

1. System Overview

This system automates the process of scraping organic Facebook posts and posts from targeted accounts, extracting content and engagement data, enriching the data with AI analysis, and storing everything in Airtable for structured access. It is designed to help companies derive their social media marketing strategy, understand target audiences, and perform R&D based on public perception.

The system is built using n8n workflows, Apify actors, and Airtable as the database and user interface.

2. Core Components

- **Airtable Base:** Acts as both the input queue for search requests and the storage for scraped results and AI-driven insights. It contains tables for **Search Requests** and **Scraped Content**.
 - **Apify Actors:** Specialized scrapers used to fetch data from Facebook.
 - **Facebook Post Scraper:** Scrapes posts, metadata (author, followers, etc.), and content URLs.
 - **Facebook Comment Scraper:** Collects comments from selected posts for deeper analysis.
 - **n8n Workflow:** The automation engine that orchestrates the entire process. It triggers on demand, calls the Apify scrapers, processes the data, sends it to the LLMs for analysis, and stores the final results in Airtable.
 - **AI Modules:** These are the intelligence layers powered by a Large Language Model (LLM) such as Gemini, which enrich the scraped data.
 - **Content Analysis:** Analyzes images and videos to identify key insights, visual hooks, undeniable proof, and overall themes.
 - **Comments Analysis:** Summarizes comment sections, identifies pain points, common questions, and determines overall sentiment.
 - **Post Review:** Provides snippets on post strengths, areas for improvement, and key learnings.
-

3. Workflow Flowchart

Airtable Search Request submitted via form →

n8n webhook triggers →

n8n picks up the request →

Apify Facebook Scraper fetches posts →

Posts are stored in the Scraped Content table →

(User can then toggle on analysis) →

Apify Comment Scraper fetches comments OR LLM analyzes content/video →

AI analysis enriches both posts & comments →

Results are available in Airtable for review.

4. Airtable Tables

4.1 Search Requests

Role: Serves as the input queue where users submit requests.

Key Fields: Search Keyword, Search Type, Number of Posts, Duration.

4.2 Scraped Content

Role: Stores all scraped Facebook post metadata and the results of the AI analysis.

Key Fields: Search Keyword, Facebook User Name, Followers, Published Date, Post URL, Views, Likes, Comments, Share count, Caption, Cover Image, Photo URL, Virality Score, Strengths, Improvement Opportunity, Lessons to Learn.

5. API & Authentication

- **Airtable API:** REST + JSON. Authentication is handled via a **Bearer Token**.
- **Apify API:** REST + JSON. Authentication is handled via an **API Token**.
- **LLM APIs (e.g., Gemini):** REST + JSON. Authentication is handled via an **API Key**.

Standard HTTP codes are used for error handling across all APIs.

6. Limitations & Error Handling

6.1 System Limitations

- **AI-Generated Data:** All AI analysis is produced by LLMs. Outputs may not always be contextually accurate, so **human review is mandatory**.

- **Scraper Dependency:** The data quality, completeness, and accuracy of parameters like dates are dependent on Apify's scraping capabilities. Changes in Facebook's structure may break the scraper.
- **Execution Limitations:** Only one webhook execution can run at a time. Each run can perform only one type of operation: Search content, Analyze video, or Analyze comments.
- **Account Search Requirement:** For account-level searches, the **exact username** must be entered. The account "display name" will not work.
- **Comments Analysis Threshold:** Comments analysis will only be triggered if a post has at least **10 comments**. Posts with fewer comments will not be analyzed.
- **Video Analysis Latency:** Video analysis relies on the Gemini LLM, which can be slow and may have a queue. While wait nodes are used, users may need to retry if no output is returned.
- **Location-Based Search:** Meta's algorithm does not share regions for posts, so location-based searches are not supported. A search for a keyword will show posts globally, not restricted to a specific country.
- **Content Type:** The system currently only scrapes and analyzes image and video posts. **Text-only posts are excluded** due to scraper limitations.

6.2 Error Handling

- **2xx:** Success.
- **4xx:** User errors (e.g., invalid request, missing fields).
- **5xx:** Server errors (retry logic in n8n is recommended).

7. Virality Score Formula

The system ranks Facebook posts by a custom Virality Score, which is a calculated field in Airtable.

Formula inputs: A post's Base Engagement Score (calculated from likes, comments, and shares), Follower Tier Multiplier, and Time Factor Score (based on recency).

This ensures high-quality, engaging content is prioritized, regardless of the creator's follower count.

8. AI Prompts (for Gemini or Similar LLM)

- **Content Analysis Prompt:** "You are given a Facebook post's caption, engagement stats, and content. Analyze the post to identify the key insights, its visual hook (what makes it a scroll stopper), any undeniable proof, and the overall theme of the post."
- **Comment Analysis Prompt:** "You are given comments from a Facebook post. Summarize the overall comments, identify common pain points, list the most frequently asked questions, and determine the overall sentiment in the comment section."

- **Post Review Prompt:** "Review the provided post. What are its strengths and areas for improvement? What are the key takeaways or lessons that can be learned from this post?"

9. n8n Workflow Structure

- **Trigger Node:** `Webhook` to initiate the workflow.
- **Airtable Nodes:** `Search` to find requests, `Create a record` to save scraped data, and `Update record` to add analysis results.
- **Apify Nodes:** `Facebook Post Scraper` and `Facebook Comment Scraper` nodes.
- **LLM Nodes:** Nodes for analyzing content and comments (e.g., `Gemini` or `GPT`).
- **HTTP Request Node:** Used for downloading media content.
- **Wait Node:** Used to enforce delays, particularly for LLM analysis.
- **Filter Node:** Used to check conditions, such as the minimum comment threshold before analysis.

10. Maintenance Checklist

- Rotate Airtable and Apify tokens regularly.
- Monitor Apify actor changes (API updates may break scrapers).
- Review n8n logs weekly for failed runs.
- Clear old requests from the Search Requests table monthly.

11. Troubleshooting

- **No posts scraped?** Check if the Facebook account is private or if the Apify actor quota has expired.
- **Comments missing?** Verify the post has at least 10 comments for analysis to trigger.
- **Rate limit errors?** Add a Wait node in n8n between Airtable writes.
- **No AI analysis output?** Check the n8n execution log for errors. The LLM may be busy; try again later.

User Guide - Detailed Walkthrough


This guide provides a step-by-step walkthrough for team members on how to use the system.

A - Access and Roles

- **Who needs access:** 1) Marketing editors (Airtable Editor), 2) Ops engineer (n8n access), 3) Data reviewer (Airtable Viewer + comments).
- **Action:** Confirm access before running searches. Ensure required Airtable permissions are set.

B - Create a Keyword Search (step-by-step)

1. Open the **Facebook search request form interface** in Airtable.
2. Populate the fields:
 - **Search Type:** Keyword Search
 - **Keyword:** [Insert Keyword]
 - **Number of Posts:** [Insert Number]
 - **Duration:** [Insert Duration]
 - **Run:** TRUE
3. **Tips:** Avoid trailing spaces in keyword fields. Use short keyword lists to reduce noise; run multiple targeted searches for variants.



The image shows a dark-themed form titled "Facebook Search Request". The form contains five input fields, each with a label above it: "Search (Keyword/Account)", "Search Type", "Max Results", "Date Range", and "Run". The "Run" field is a checkbox. All input fields are currently empty.

C - Create an Account Search (step-by-step)

1. Choose **Account Search** in the Search Type field.
2. Enter the exact Facebook username (e.g., PageName) in the Facebook Account field. Do not use the display name.
3. Set the Number of Posts and Run to TRUE.

- **Action:** Check Apify run details and the dataset to verify returned items.

Facebook Comments Scraper - Run

✓ Succeeded

Finished! Total 3 requests: 3 succeeded, 0 failed.

RESULTS	REQUESTS	PRICE ⓘ	STARTED	DURATION
12	3 of 3 handled	\$0.029	2025-08-29 23:53	10 s

ⓘ

Help Apify improve your experience with Facebook Comments Scraper

We need your permission to share all past and future runs of this Actor with the community. You can revoke your consent anytime in [Settings > Privacy](#).

Yes, I agree

No, thanks

Output 12

Log

Input

Storage

Live view

Triggered integrations 0

Overview

All fields

#	Post author postTitle	Comment text
1	Most robots today work on their own. They make decisions without understanding what other robots and humans are doing in the same space...	Nokia still exists

F - Review Scraped Content

1. Open the **Scraped Content** table or the `Scraped Content` interface in Airtable.
2. Open a record to confirm the data: Caption, Views, Likes, Comments, Shares, etc.
3. Look for the `Analyze Content` and `Analyze Comments` toggles.

4. **Reviewer tips:** If a field is blank, check if the data was available for that post. Compare the output to the source post in the Apify dataset if data looks missing or odd.

Facebook Post Scraper System - Keyword - Scraped Content						
Search Keyword	Search Type	Facebook User Name	Post URL	Post Type	Published Date	Virality Score
Cetaphil	Keyword	American Beauty BD	https://www.facebook.co...	post	2025-08-27	21
Cetaphil	Keyword	Sanjida Mun Appi	https://www.facebook.co...	post	2025-08-27	24
Cetaphil	Keyword	Cetaphil Myanmar	https://www.facebook.co...	post	2025-08-27	25
Cetaphil	Keyword	Philstar.com	https://www.facebook.co...	post	2025-08-27	35
Cetaphil	Keyword	Sanjida Mun Appi	https://www.facebook.co...	post	2025-08-27	24
Ikea	Keyword	Norwich Evening News	https://www.facebook.co...	video	2025-08-28	43
Ikea	Keyword	Tia giftshop penkridge	https://www.facebook.co...	post	2025-08-28	19
Nokia	Keyword	Ahmed The Nature Lover	https://www.facebook.co...	video	2025-08-28	37
Nokia	Keyword	Meu Smart	https://www.facebook.co...	post	2025-08-28	27
Nokia	Keyword	Price Drop	https://www.facebook.co...	post	2025-08-28	28

G - Request Analysis (content/comments)

1. Toggle **Analyze Content** or **Analyze Comments** to **TRUE** on a record to trigger the analysis.
2. **Note:** For comments analysis, the post must have at least 10 comments.
3. **Action:** Monitor n8n runs for AI processing completion.
4. **Timing guidance:** LLM jobs may be queued; expect variable completion times depending on server load.

Facebook Post Scraper System - Keyword - Scraped Content										
Score	Followers	Comments	Likes	Views	Share count	Caption	Cover Image	Duration	Analyze_Content	Analyze_Comment
	1,046,810	6	37	0	0.0	🔥 Customers Favorite glo...		0:00	✓	✓
	13,407	18	10	0	1.0	আমাদের বাসায় একদিন...		0:00	✓	✓
	158,051	8	59	0	1.0	All sold out 🛒🔥🔥🔥...		0:00	✓	✓
	2,986,750	11	1,024	0	18.0	Say goodbye to redness a...		0:00	✓	✓
	13,407	18	10	0	1.0	আমাদের বাসায় একদিন...		0:00	✓	✓
	81,496	364	510	47,138	70.0	IKEA has opened its doors!...		0:16	✓	✓
	972	4	1	0	1.0	Morning everyone 🌞. We...		0:00	✓	✓
	49,481	21	545	6,332	4.0	Nokia 17 Pro Max 🌟🔥		0:17	✓	✓
	639,617	7	77	0	7.0	Blackberry VS Nokia		0:00	✓	✓

H - Interpret AI Outputs and Take Action

Key fields produced by AI:

- **Virality Score:** Use to prioritize content.

- **Key Insights:** Summary of the post's value.
- **Visual Hook:** What makes the post a "scroll stopper."
- **Undeniable Proof:** Credible elements of the post.
- **Comments_Sentiment / Comments_Summary:** Insights from the comment section.
- **Post Review:** Strengths, weaknesses, and key takeaways.
- **Suggested next steps based on outputs:**
 - High Virality + positive insights: Add to a content replication pipeline or creative brief.
 - Negative sentiment: Route to product or support teams.
 - High number of questions: Collect into an FAQ.

stem > Keyword - Content Analysis

h Type	Facebook User Name	Post Type
keyword	Philstar.com	post
keyword	Sanjida Mun Appi	post
keyword	Norwich Evening News	video
keyword	Tia giftshop penkridge	post
keyword	Ahmed The Nature Lover	video

^

v

🖨

🔗

Ikea

Search Type

Keyword

Facebook User Name

Norwich Evening News

Post Type

video

Post URL

<https://www.facebook.com/reel/1305063231121635/>

Key Insights

The video announces the official opening of a new IKEA store, marked by the iconic blue and yellow facade, a celebratory balloon arch, and staff in branded shirts. It then powerfully showcases the significant public anticipation and demand by panning across an incredibly long queue of eager customers stretching far down the side of the building, all waiting to enter the newly opened store at Hall Road Retail Park.

Visual Hook

The immediate, bright shot of the iconic IKEA store with its large blue and yellow logo and festive balloon arch, coupled with the clear "IKEA has officially opened!" text, instantly captures attention by signalling a highly anticipated and recognizable retail event.

Undeniable Proof

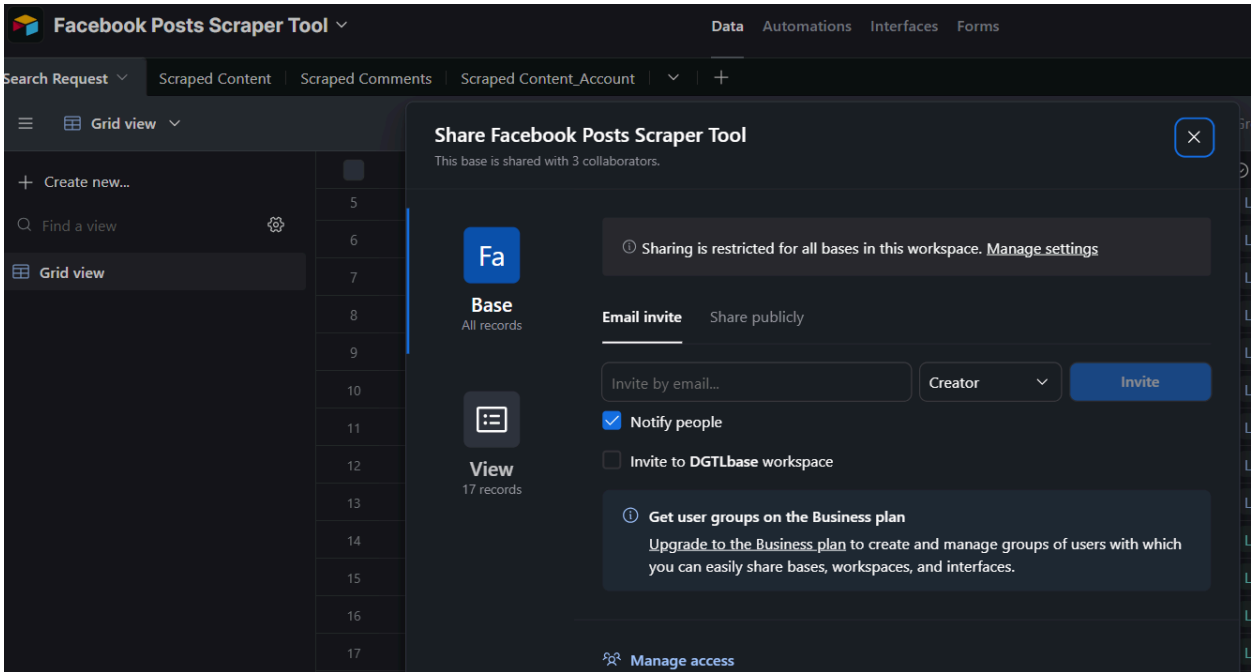
The exceptionally long and continuous queue of people extending for hundreds of feet along the building's exterior provides powerful social proof, demonstrating widespread public

Keyword - Comments Analysis			
Facebook User Name	Post Type		
Philstar.com	post		
Sanjida Mun Appi	post		
Norwich Evening News	video		
Tia giftshop penkridge	post		
Ahmed The Nature Lover	video		

Cetaphil	
Search Type	Keyword
Facebook User Name	Sanjida Mun Appi
Post URL	https://www.facebook.com/sanjida.mun.appi/posts/pfbid02npbP hhtJ8qXQuwCtJ2aktoXQPEoLPecob4BuFM4M2WTCdvUu57LtNLu SJ2w5p5UJ
Post Type	post
Published Date	2025-08-27
Virality Score	24
Comments	18
Comments Summary	Audience enjoyed the humorous, relatable family scenario featuring Cetaphil, while debating foam/lather, real-life effectiveness, and product suitability. Many praised the gentle nature but remained curious about non-foaming performance, availability, and whether face wash can double as handwash.
FAQ in Comments	<ol style="list-style-type: none"> 1. Why does Cetaphil produce less foam or different bubble sizes? 2. Does less foam still clean effectively? 3. Is Cetaphil suitable for sensitive, oily, or acne-prone skin?

I - Exporting / Sharing

You can export the data directly from Airtable using its built-in export features.



Appendix

- **n8n Workflow JSON (full export):**
<https://dgtlbase.app.n8n.cloud/workflow/OG06qqt0SVuq4ke2>
- **Screenshots of Airtable schema:** <https://airtable.com/appbKMler8cGNdJm0/api/docs>
- **Prompt library for all AI modules:** See Section 8 and the n8n workflow.
- **Sample JS snippets for webhook automation within Airtable:**

JavaScript

```
let table = base.getTable("Search Requests");
let record = await table.selectRecordAsync(recordId);

await fetch("YOUR_N8N_WEBHOOK_URL_HERE", {
  method: "POST",
  headers: { "Content-Type": "application/json" },
  body: JSON.stringify(record)
});

console.log("Sent to n8n:", record);
```