

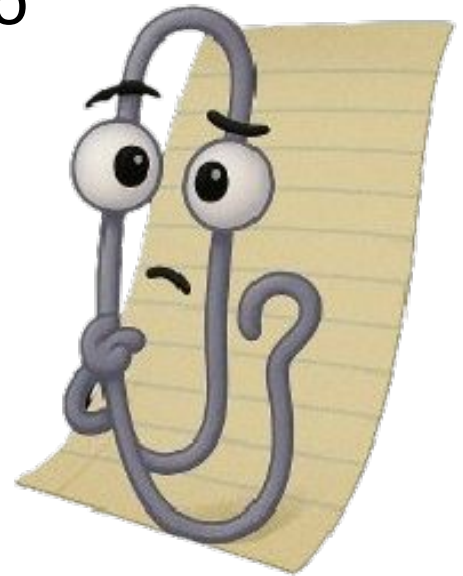
# SocialOSINTLM

Fusing Multi-Platform Data with  
Advanced Language & Vision Models

Brenden Miller

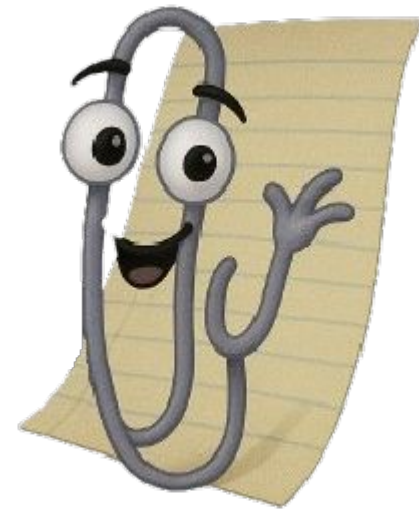
# Origin story

- What does my online persona say about me
- Spiralled to:
  - What can we find with the open info



# What is SocialOSINTLM?

- It's just an aggregator and analyser.
- Using OpenAI compatible API for use with hosted and local models
- **Twitter**
- **Bluesky**
- **Reddit**
- **Mastodon**
- **Hacker News**
- --offline available cache limit bypass



# The Power of Fused Intelligence

- SocialOSINTLM's power comes from fusing text and visual data.

## Cross-Platform Understanding:

- Actionable Output: The result is a structured analytical report, available in Markdown or JSON, tailored to your specific query and providing a clear summary of findings.

# The Intelligence Engine: Data Ingestion & AI Core

- Visual Context Capture: Downloads publicly shared images for analysis by AI vision models.

## Grounded Text and Image prompts

- Behavioural patterns (e.g. engagement styles, platform-specific activity)
- Key semantic themes, recurring topics, and sentiment.
- Communication styles and linguistic nuances.
- Objective Vision Analysis: Shared images are interpreted by vision-enabled LLMs for objective scene description, focusing on:
  - Settings, environments, and potential geographic indicators.
  - Key objects, legible text within images, and distinct symbols.
  - Apparent activities or events depicted.
- This process strictly avoids PII speculation, concentrating on verifiable visual evidence to add context.

# Image analysis - Example

"analysis": "Here is an objective OSINT analysis of the image:\n\n\* \*\*Setting/Environment:\*\* Appears to be inside a mode of transportation, likely an airplane or potentially a train, based on the visible tray table, seat back pocket with a magazine labeled \"Bistro,\" and what appears to be a window or seat back in the background. The lighting suggests daytime.\n\n\* \*\*Key Objects/Items:\*\*\n\n\* Knitting needles with a small swatch of multi-colored yarn.\n\n\* A ruler with knitting and crochet needle sizes labeled in US and mm sizes. The ruler is labeled \"ChiaoGoo\" and \"Premium Knitting & Crochet Too...\"\n\n\* A sheet of paper with knitting instructions.\n\n\* Sunglasses resting in a seat back pocket.\n\n\* A magazine with \"Bistro\" visible on the cover or spine, also in the seat back pocket.\n\n\* A blue clip or binder clip visible above the magazine in the seat back pocket.\n\n\* \*\*People (if present):\*\* No people are visible in the image.\n\n\* \*\*Text/Symbols:\*\*\n\n\* \"Bistro\" (on the magazine).\n\n\* \"ChiaoGoo\" (on the ruler/tool).\n\n\* \"Premium Knitting & Crochet Too...\" (on the ruler/tool).\n\n\* \"US\", \"mm\", and various numerical measurements and abbreviations (e.g., K1, P1, st, cm, RS, WS) on the ruler and instruction sheet.\n\n\* \"Neckband (photo support on pages 9 and 10)\" (on the instruction sheet).\n\n\* \"Right Side\" (on the instruction sheet).\n\n\* Instructions detailing knitting steps (e.g., \"With 4 mm needles, CO 11 sts using the long tail\", \"Row 1 (RS) - K1, \*P1, K1\*\", \"Repeat Rows 1 and 2 until work measu...\", \"(16) 17 (17) 17 cm.\", \"End by working a\", \"let sts rest.\", \"Left Side\", \"With RS facing, pick up 11 st\", \"cast on with a new piece of\", \"edge to left.\", \"The yarn you\", \"the needle's opposite en\", \"Row 1 (RS) - Slip 1 w\").\n\n\* \*\*Activity/Event:\*\* The image depicts the activity of knitting, specifically working on a small swatch, likely following instructions on the adjacent paper.\n\n\* \*\*Implicit Context Indicators:\*\* The tray table, seat back pocket, and magazine labeled \"Bistro\" strongly suggest a setting involving travel, likely commercial air or rail transit. The multi-colored yarn suggests a colorful or creative project. The detailed knitting instructions indicate focused work on a specific pattern.\n\n\* \*\*Overall Scene Impression:\*\* A focused activity (knitting) taking place in a travel environment, likely an airplane or train. The scene suggests a leisure or productive activity during transit.",

"url": "https://cdn./img/...@jpeg",

"alt\_text": "Beginning of the knit project on a tray table.",

Image from post

## SocialOSINTLM

Collects and analyzes user activity across multiple platforms using vision a  
Ensure API keys and identifiers are set in your `.env` file.

### Select Platform(s) for Analysis:

1. Twitter
2. Bluesky
3. Mastodon
4. Reddit
5. Hackernews
6. Cross platform
7. Purge data
8. Exit

Enter number(s) (e.g., 1 or 1,3 or 6) (8): |



# Account comparison

Analysis Query> (): Are these two accounts by the same person/company?

Starting analysis for query: 'Are these two accounts by the same person/company?'

## Analysis Report

### OSINT Analysis Report

Query: Are these two accounts by the same person/company? Targets Queried: bluesky, twitter Report Generated: 2025-05-19 11:54:  
UTC Mode: Online Models Used:

- Text Analysis: [google/gemini-2.0-flash-001](#)
- Image Analysis: [google/gemini-2.0-flash-001](#)

Based on the provided data, it is highly likely that the two accounts, [@FT](#) on Twitter and [financialtimes.com](#) on Bluesky, represent the same entity: the Financial Times newspaper.

#### Supporting Evidence:

- **Profile Information:**
  - The Twitter account [@FT](#) has a description that links to the Financial Times website (<https://t.co/EYmAcRLBHv>) and encourages registration for free articles. The location is listed as London.
  - The Bluesky account [financialtimes.com](#) has a display name of "Financial Times" and a bio that states, "The best of FT journalism, including breaking news and analysis," with a link to the FT website ([www.ft.com](http://www.ft.com)). It also states that the users the account follows are verified FT staff or contributors.
- **Content Similarity:** Both accounts consistently post links to articles on the Financial Times website. The headlines and topics covered are the same across both platforms (e.g., "UK and EU agree post-Brexit reset," "US Treasury yields reach 5%," "Far right surges in Portugal").

# Account mismatch

```
Enter number(s) (e.g., 1 or 1,3 or 6) (8): 1,2
Selected: Twitter, Bluesky
Enter Twitter username(s) (comma-separated, no '@')) (): ft
Enter Bluesky username(s) (comma-separated, e.g., 'handle.bsky.social')) (): theonion.com
Initializing API clients...
```

## Analysis Session

```
Targets: Twitter: @ft | Bluesky: theonion.com
Enter your analysis query below.
Commands: 'exit' (return to menu), 'refresh' (force full data fetch), 'help'
```

```
Analysis Query> (): Are these two accounts by the same person/company?
```

```
Starting analysis for query: 'Are these two accounts by the same person/company?'
```

## Analysis Report

### OSINT Analysis Report

```
Query: Are these two accounts by the same person/company? Targets Queried: bluesky, twitter Report Generated: 2025-0
UTC Mode: Online Models Used:
```

- Text Analysis: `google/gemini-2.0-flash-001`
- Image Analysis: `google/gemini-2.0-flash-001`

Based on the provided data, it is highly unlikely that the @FT Twitter account and the theonion.com Bluesky account by the same entity.

### Evidence & Analysis:

#### 1 Semantic Content & Themes:

- @FT (Financial Times) consistently posts news articles and updates on business, finance, economics, and international affairs. The tone is generally objective and informative.
- theonion.com (The Onion) exclusively posts satirical articles and content. The tone is humorous and often absurd.
- The subject matter and tone are diametrically opposed.

# OSINT Practitioner

Starting a new job

Acquiring a company

Missing persons

Relationship matching

Red team - recon

Information Leakage

# Applications & Benefits for Technical Professionals

- Interactive CLI with rich formatting for ease of use.
- Input via JSON from stdin.

```
"platforms": {  
  "twitter": ["twitterhandle"],  
  "reddit": ["reddithandle", "reddithandle2"],  
  "hackernews": ["hackernewshandle"],  
  "bluesky": ["handle.bsky.social", "handle2.bsky.social"],  
  "mastodon": ["handle@mastodon.social"]  
},  
"query": "Summarize main topics discussed and analyze sentiment",  
"format": "markdown"
```

- Compliance & Respect:
- 24 hour cache and 'Purge Data' option. This respects the user's 'right to be forgotten' principles, allowing you to clear cached data and media.

# Technical Foundations & Ethical Framework

- Exclusively Public Data: Operates strictly on publicly available information accessed via platform APIs.
- Data Control & Minimization: Focuses on recent activity and incorporates a 'Purge Data' functionality to delete all cached data, media, and reports, respecting data privacy principles.

- Acknowledge model bias, stay in the loop

Thank you

<https://github.com/bm-github>