

# EVERDIAN

Stream AI Framework for Real-Time Insights

# The Haystack Challenge

When critical signals are buried in massive and fragmented information flows, forcing organizations into slow and reactive decision-making.



## Incomplete and Unstructured Information

Multilingual, multimodal data remains messy and disconnected, obscuring the signals that matter.



## Overlooked Risks and Missed Opportunities

When data is poorly visualized and trapped in silos, critical risks are missed and opportunities slip away.



## Delayed Awareness

Manual, fragmented collection makes it impossible to monitor all relevant sources rapidly and at scale.

# Platform

With over ten years of industry expertise in developing tailored AI models and data ingestion pipelines at scale, Everdian's Stream AI turns sweeping volumes of data from diverse digital ecosystems into real-time, actionable insights.



Stream AI is... In Real Time, Multimodal, For all Datasets, at Scale

Confidential

20251103-AL

# Our Edge

Unique platform advantages, driven by a world-class team

## Massive Multimodal Ingestion

Process structured and unstructured data in over 200 languages, including text, images, video, audio, and live streams.

## Modular & Extensible

AI bundles and customizable widgets adapt seamlessly to industry requirements and workflows.

## Real-Time at Scale

Analyze millions of events with low latency, enabling mission-critical decision-making.

## Continuously Evolving

A scalable architecture that grows with your data and unlocks new AI capabilities on shared or dedicated infrastructure.



# Illustrative Use Cases

Applications powered by the AI Engine include :



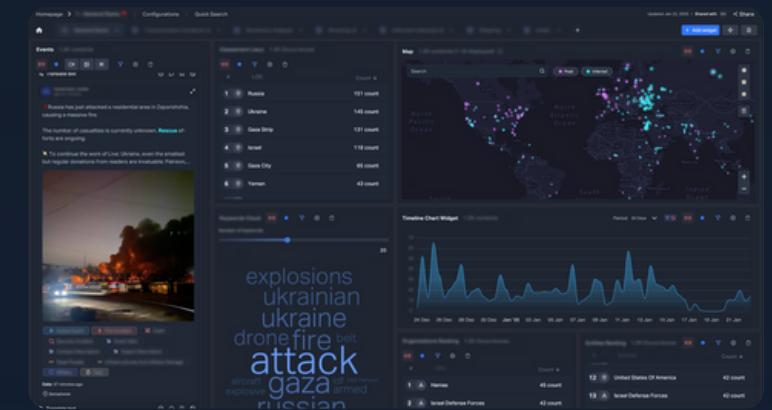
## Disinformation Detection & Monitoring



Uncover, analyze, and visualize malicious narratives confidently in real-time.



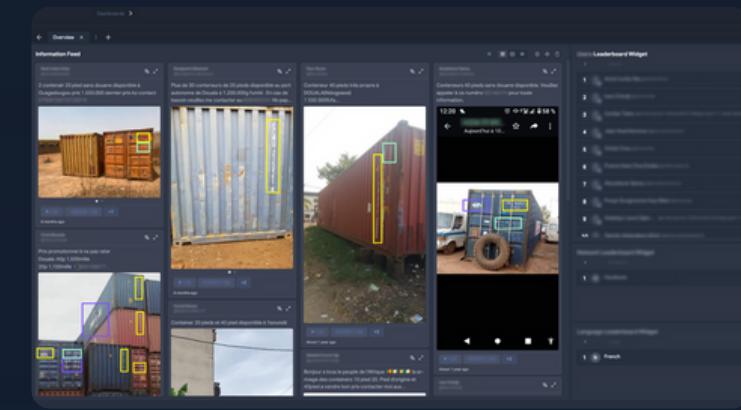
## Real-Time Security Alerts & Situational Awareness



Detect global threats in real-time.



## Fraud Prevention, Detection & Investigation



AI-powered insights to uncover, investigate, and disrupt fraudulent activity.



## Critical Insights & Strategic Decision-Making



## Risk Management & Mitigation



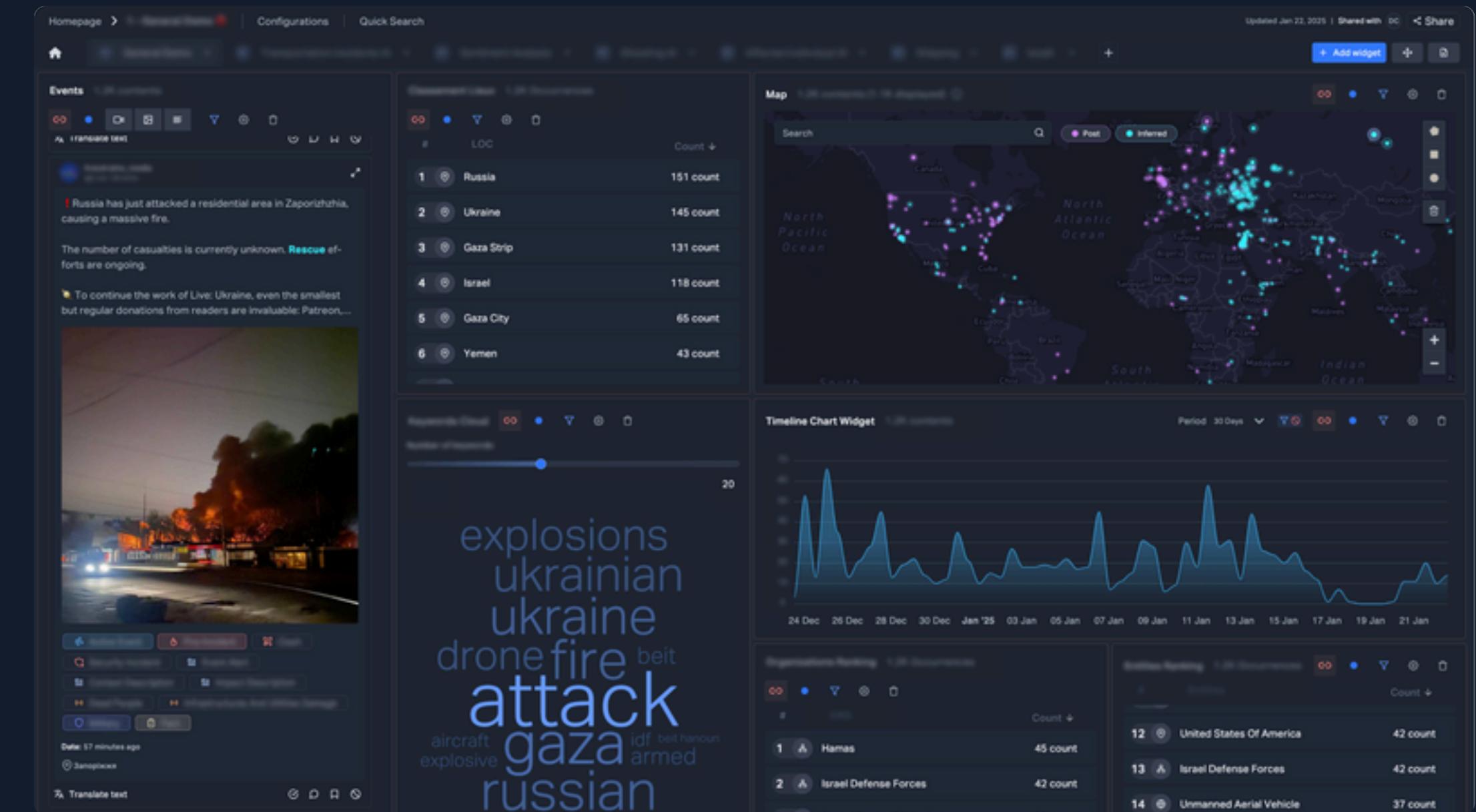
## Proactive Digital Content Moderation

# Real-Time Security Alerts & Situational Awareness

Stay ahead of evolving threats with AI-powered monitoring and real-time risk visualization.

## Detect global threats in real-time

Gain a critical time advantage through immediate threat identification and comprehensive risk visualization. With expansive data scraping, 50+ AI Models, and capacity for 200+ languages, Everdian provides the foresight needed to detect risks across truly global operations and take decisive action against them.



# Detecting and Monitoring Disinformation

AI-powered intelligence to uncover, analyze, and counter malicious narratives.

## Combat Disinformation with Precision

Uncover, analyze, and visualize malicious narratives confidently in real-time. Our platform's AI models meticulously sift through vast volumes of data in real-time to provide you with the actionable insights needed to safeguard the integrity of your information ecosystem, operations, and public reputation.



# Industries We Empower

Our AI Engine adapts to diverse industries, powering mission-critical applications:



Government  
& Public Sector



Financial Services



News & Media



First Responders  
& Law Enforcement



Enterprises &  
Brands



NGOs & Humanitarian



Defense



And More...

# Business Model

Recurring subscriptions, usage-based tokens, and professional services

## Licence Fees

### Annual Subscription

Access to the platform.



Data  
Collection

### Tokens

Flexible, consumption-based pricing model.  
Breakdown of token consumption :



Data analysis  
with AI



Data Storage,  
Curation and Type

### Training

Hands-on training  
with our team.

### IT Support

Dedicated support  
to ensure seamless  
operations.

### Consulting

Ensuring effective  
platform use for your  
mission.

### Development

Custom projects to  
meet specific needs.

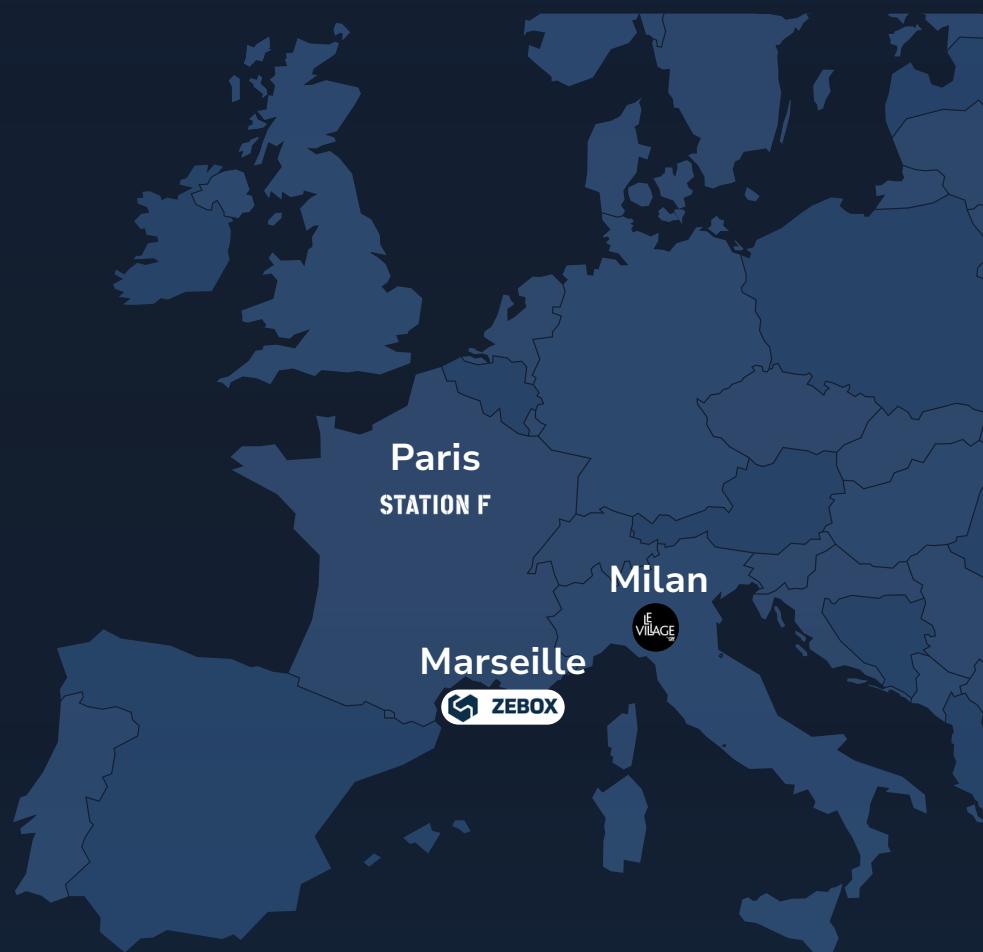
# About Us

Proven company, trusted collaborations, recognized programs

An award-winning start-up in the world's most prestigious innovation programs



## Offices



## Management



Quentin Lhomme  
Chief Executive Officer



Cédric Milinaire  
Chief Technology Officer



Clélia Chollon  
Head of Operations



Jeremy Auguste  
Head of AI - PhD



Blandine Berthod  
Product Manager



# EVERDIAN

Everdian x Albert School Project

Exploring Real-Time Event Intelligence with Natural-Language Queries

# What Clients Can Do Currently

The screenshot displays the EVERDIAN Stream AI platform's user interface. On the left, there is a sidebar titled "Event detection demonstration" with various filtering options: Date (from 01/10/2023 to 12/10/2023), Interactions, Keywords, Sources, and Users. Below this is a "Classification Et Détection D'événements" section with categories like Accident aérien, Accident de la circulation, and Conférence. A search bar at the bottom shows "Recherche" and "449 résultats". The main area features a map titled "Map - 449 contenus (449 affichés)" showing a concentration of events in South Asia, specifically in Pakistan, India, and parts of Central Asia. To the right of the map are several analytical widgets: "Classement Lieux - 449 Occurrences" listing locations like Pakistan (162 posts), Lahore (49 posts), and Balochistan (34 posts); a "Timeline Chart Widget - 449 contenus" showing a fluctuating timeline from Oct 02 to Oct 29; "Organisations Ranking - 449 Occurrences" listing organizations such as Tehrik-e-Talib... (57 posts) and North Waziristan (26 posts); and "Entities Ranking - 449 Occurrences" listing entities like Pakistan Army (29 posts) and Pakistan (142 posts).

## Filtering capabilities:

- Date ranges
- Geolocations (free-form areas, administrative areas)
- Sources
- Event types
- Keywords
- Disambiguated Entities
- Users

## Other capabilities:

- Data exports
- Event deduplication and clustering
- Automatic translation (200 languages)
- Automatic transcription (100 languages)

# What Clients Would Like To Do

Exploring Real-Time Event Intelligence with Natural-Language Queries

## Non-technical Users

- Don't want to think about how the platform works
- Ideally, only want to ask the platform in natural language what they want in order to find specific kind of events
- Output must be in the form of reports, tables, charts and other forms of pre-digested data.

## Power Users

- Want to keep access to the raw data
- Want to expand on available filters
  - Use natural language to define new type of unsupported or very complex filters
- Want to apply complex operations on the data directly in the platform
- Want to question the data itself

**Core challenge:** Enable any user to explore millions of contents through simple, conversational queries.

# Current limitations

## Filter complexity

Manual chaining of precise filters (location + event type + time) and widgets creates steep learning curves for non-experts

## Inflexible exploration

Complex queries like "Show verified port intrusions excluding false alarms" outpace current tools

## Noise in results

ML/DL model errors and rigid filtering can yield false positives (e.g., misclassified "drone incidents")

## Unsupported analysis types

The platform lacks native capabilities for some critical analytical workflows required by analysts, necessitating manual workarounds

**Result:** Analysts spend excessive time on data exploration and manual work rather than threat response, with key insights obscured in raw event streams.

# Solution Requirements

**Enable platform users to explore and interact with the prequalified event database using simple, conversational queries.**

The new solution should allow the following features :

- Generating detailed situational reports
- Producing aggregated statistics and visualizations (graphs, tables, summaries)
- Identifying how events are being discussed or framed
- Locating relevant, specific content quickly
- Detecting recurring event patterns
- Interacting with external tools and resources
- ...

# Industrial constraints



## Multilingual and heterogeneous data

Collected contents can be in 200+ languages, and our platform itself is used by users that speak different languages.

Contents can be ingested from a wide variety of sources.



## Massive quantities of data

The platform handles millions of contents per day. As such, technical solutions need to be economically sound.



## Sensitive use-cases

Multiple clients have critical use-cases:

- Analysis should minimize hallucinations as much as possible.
- Models that are used in the end product need to be deployable on any infrastructure.

# Provided Database

Sample:

```
{  
    "id": "d48ryhz8j6yg0tqysc82ikf",  
    "text": "Revieve Launches AI-Powered Skin Age Diagnostics, with RoC Skincare as First Adopter",  
    "english_sentence": "Revieve Launches AI-Powered Skin Age Diagnostics, with RoC Skincare as First Adopter",  
    "lang": "eng_Latin",  
    "labels_v2": [  
        {  
            "type": "Event Temporality",  
            "value": "Active event",  
            "score": 0.9999768733978271  
        },  
        {  
            "type": "Main Categories",  
            "value": "Other",  
            "score": 0.9538894891738892  
        },  
        {  
            "type": "Content Type",  
            "value": "Context description",  
            "score": 0.9395267963409424  
        },  
        { "type": "Statement-Type", "value": "Fact", "score": 0.999994158744812 }  
    ],  
    "publish_date": "2025-08-06T23:59:51.000Z",  
    "locations": { "mentions": [], "inferred": [] },  
    "images": [  
        "https://mma.prnewswire.com/media/2744873/ReviewexRoCSkincare.jpg?p=facebook"  
    ],  
    "videos": [],  
    "url": "https://news.knowledia.com/all/all/articles/revieve-launches-ai-powered-skin-age-diagnostics-with-ro",  
    "user": {  
        "userName": "news.knowledia.com",  
        "metrics": [{ "metricName": "Rank", "metricCount": 124047 }]  
    },  
    "network": "news"  
}
```

## Characteristics

- 4M+ contents from social networks and online press
- Collected from a general configuration used for client demonstration
- Original collection based on generic keywords, no specific type of event was targeted
- From 2025-08-06 to 2025-10-27, with missing data for some dates
- Pre-qualified: only contents that talk about “events” are kept
- Contents are split by day, and stored in the JSONL format
  - JSONL → One JSON per line

A **schema.json** file is available with the data in order to precisely describe each field.

# Provided Database

What can be found in the database

## Contents

- Text:** The text in the original language associated with the collected content
- English Sentence:** The text translated to English
- Lang:** The language of the original text

## Geolocations

- Inferred:** locations of the event, inferred automatically
- Mentions:** locations that are simply mentioned in the text.

## Images and videos

Some contents may include images or videos. These are represented by their URLs.

## Users

Contains information about the author of the post such as the username and metrics such as number of followers

# Provided Database

## Main Categories

- Airplane Accident
- Roadway Accident
- Railway Accident
- Nautical Accident
- Illicit / Illegal Action
- Arrests
- Unidentified Attack
- Other
- Shelling / Rocket Incident
- Natural disasters
- Clash
- Power Outage
- Cyberattack
- Hazard
- Structure Collapse
- Kidnapping
- Criminal Investigation
- Explosion
- Fire Incident
- Security Incident
- Non-Firearms Weapon Incident (bladed, blunt)
- Judgement/Sentence
- Maritime related
- Lawmaking
- Protest
- Piracy
- Suing
- Shooting
- Suspecting
- Drug Trafficking

## Impact

- Arrested People
- Missing People
- Infrastructures and Utilities Damage
- Dead People
- Displacements and Evacuations
- Affected Individual
- Injured People
- Found People
- Condemned People
- Hostage

## Temporality

- Active event
- Past event
- Future event

## Statement Type

- Fact
- Information
- Opinion

## Content Type

- Event alert
- Analysis/Statistics
- Context description
- Impact description
- Hypothetical/Fictional
- Reference to event

## Law Enforcement Agency

- Police
- Military
- State Forces - Other

# Provided technical tools

## Access to Scaleway Generative APIs

- Documentation: <https://www.scaleway.com/en/docs/generative-apis/quickstart/>
- List of supported models: <https://www.scaleway.com/en/docs/generative-apis/reference-content/supported-models/>

## Recommended models

- mistral-small-3.2-24b-instruct-2506
- qwen3-235b-a22b-instruct-2507



Each team will have access to its own API Key.

Usage is monitored, and access will be revoked in case of unreasonable usages.

# Main Expectations

## Focus on quality of the results

We'd rather have a few well implemented features from the original requirements.

## Works with our technical stack

The platform handles millions of contents per day. As such, technical solutions need to be economically sound.

## Works with all quantities of data

Make sure that the solution works with the full dataset but also with a limited quantity of data.  
Don't hesitate to add your own filters to the data since power users may use the platform that way.

## Multilinguality

The solution should be able to produce multilingual outputs, with a focus on English and French

# Example Scenario 1

## User query

**"Give me a global report about Gaza situation from today"**

## Possible output result

**General Overview :** The Gaza Strip is entering its 389th day of humanitarian crisis since the beginning of the conflict. Despite a partial ceasefire in effect for two weeks, tensions continue to flare in several areas of the southern territory.

Timeline of the Day

### **Detected Event :**

- 6:30 a.m. – Rafah : Israeli drones flew over the southern border for nearly an hour, with no confirmed strikes. However, residents reported warning shots near a humanitarian convoy attempting to reach the Kerem Shalom crossing.
- 12:00 p.m. – Gaza City : A small group of young engineers restarted a makeshift solar station to power a field hospital. The local NGO Nafas estimated that this would allow respirators to operate for an additional eight hours each day.
- 1:45 p.m. – Political Statement : The Israeli Prime Minister declared during a press conference in Tel Aviv that “any Palestinian armed presence beyond Wadi Gaza will be considered a threat.” Hamas responded with a statement condemning “a provocation that violates the ceasefire.”
- 8:00 p.m. – Nuseirat Camp : An accidental fire destroyed twenty tents in a displacement camp. Volunteer firefighters contained the blaze after two hours. Three people were killed and fourteen injured. The fire is believed to have started from a candle used during a power outage.
- 10:30 p.m. – Rafah : Discreet talks were held between a Qatari envoy and representatives of several Palestinian factions regarding a proposed 72-hour humanitarian truce to allow repairs on the main desalination plant. No public agreement has yet been announced.

# Example Scenario 2

## User query

"I want a tabular report about yesterday's road accidents in Saint-Étienne"

## Possible output result

Date	Location	Type of Accident	Cars	Pedestrian	Motorcycle	Bus	Minor Injuries	Serious Injuries	Fatalities	Road Conditions
21/10/25	Boulevard Thiers	Rear-end collision	2	0	At least one	0	1	N/D	N/D	Dry
21/10/25	Rue Bergson	Pedestrian hit on crosswalk	1	1	0	0	x	1	0	N/D
21/10/25	Avenue Denfert-Rochereau	Multi-vehicle pile-up	3	N/D	0	1	5	1	1	Wet
21/10/25	Tunnel de la Terrasse	Single-vehicle accident	N/D	N/D	1	N/D	0	1	2	Dry

# Other types of valid queries

These are only examples, please experiment with many more queries !

## User query

"Récupère-moi les contenus les plus révélateurs de la situation à Marseille cette dernière semaine."

## User query

"Identify the main narratives surrounding the national guard deployment in Memphis"

## User query

"Export the current dataset to Excel and create a dashboard summary."

## User query

"Take our alerts dataset, pull nearest-station weather metrics (temperature, precipitation, humidity) from an external weather DB for the most recent accident alerts, compute aggregated statistics by weather condition, and produce a CSV report ."

## Winner Prize

- **1 internship at Everdian**
  - with the R&D team and the Business team
  - in Marseille (Zebox - Grand Central) or Paris (Station F)
  - 4-6 months
- **3 other prizes among (non-exhaustive):**
  - Delta Festival tickets
  - OM Match tickets
  - Mistral API credits / Mistral Pro subscription



# Contact

@ jeremy.auguste@everdian.com

Everdian.com