
Actionable Intelligence

Our Software scans billions of data points every day, detecting both physical and narrative threats in real time and providing organizations with the context they need to make mission critical decisions.

Detection as a Service

DaaS White Paper

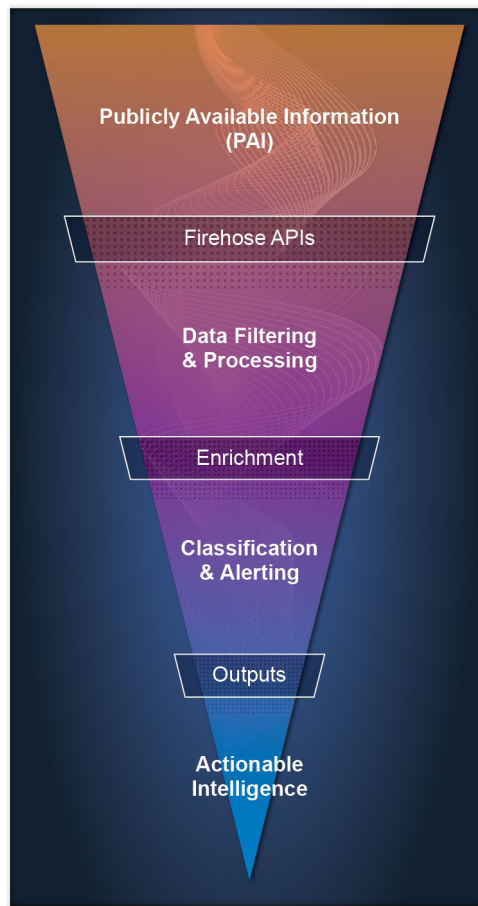
July 2022

Point of Contact:

Byron Carroll, President
Carroll Woods Inc. D.B.A Carroll International
bcarroll@carrollcommunications.guru
+1 (910) 653-2386
www.carrollintl.com

DaaS White Paper

July 2022



Executive Summary

This platform provides powerful indications and warnings, measures of effect, and analysis, enhancing situational understanding of both the cyber and physical environments. The holistic view of the digital media landscape combined with our automated analytics, allows analysts to easily identify and separate specific cyber influencers from the flood of informational data. Monitoring these individuals or groups establishes patterns and techniques used to manipulate the conversation surrounding local (and global) events. Integrating this data into the broader operational picture provides a comprehensive understanding of how specific influences and/or centers of gravity can affect friendly or malign efforts.

Through direct data provider relationships, this platform ingests billions of raw data and metadata points from over 50 million unique sources by leveraging multiple PAI full Firehose APIs. Most importantly, the platform has full Firehose API access to valuable and relevant data feeds including Sina Weibo (SW), VKontakte (VK), and Twitter. These social media platforms represent the largest social networking platforms in China, Russia, and the wider world, respectively. This level of data access guarantees nothing is missed in the information environment, while preserving valuable metadata, allowing the platform to employ advanced analytics and proprietary AI/ML enrichment.

Technology Concept

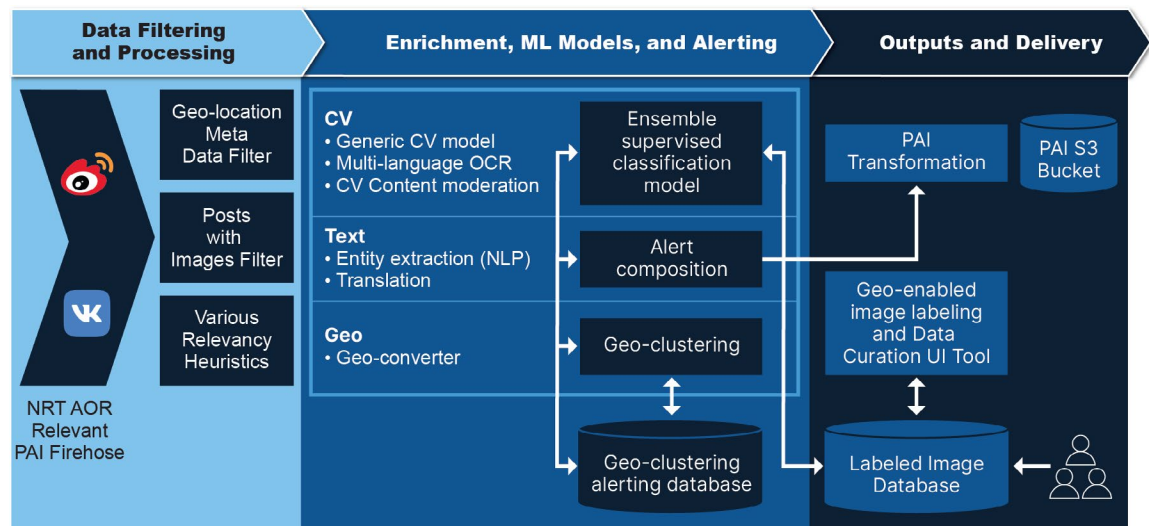
The ingestion model isolates and enriches data received with Exchangeable Image Files (EXIF), using proprietary AI/ML models to enable CV and NLP processing to assign pre-specified weighted labels and detect narratives (individual data points) for Areas of Interest (AOIs). Indications and Warning (I&W) alerts are triggered in NRT from a rules-based heuristic scoring model, informing users of suspicious activity, malign narratives, geographic sources, and providing actionable insights. Detections as a Service (DaaS) merges CV, NLP, Exif data, and integrates a narrative intelligence contextual layer. I&W alerts are delivered in JSON format along with the raw data and associated media files.

Data is ingested via full Firehose access to Russian (VK) and Chinese (SW) social media data and run through two filters, first for any post that contains imagery and second for posts with imagery that contain geo-located or geo-referenced metadata, leaving only mission-relevant imagery. Heuristics are then applied to the remaining

data containing only mission-relevant geo-located/geo-referenced posts with imagery, including searches for relevant keywords and geographic bounding box exclusions (avoiding false positives). For example, there are ~15M daily posts in the EUCOM AOR on VK. Of these ~15M posts, 500,000 contain posts with imagery and geo-located metadata. Mission-relevant imagery will then be derived from the 500,000 remaining posts. Below is an example of our Data Filtering & Processing model.

Approach

The platform examines specially curated data feeds made available via our access to full Firehose APIs and taxonomy research. Using a combination of proprietary algorithms and imagery interpretation software, DaaS identifies and isolates visual media (photos and videos) depicting military personnel, equipment, vehicles, and installations. Once media is detected, the platform analyzes its recency and impact on the information environment. Media impact is determined by examining environmental baselines and observing if there is a heightened or abnormally high interest in the imagery. Measuring impact against a baseline provides a real-time, continuously evolving data set with immediate utility and applicability to stakeholders interested in developing OSINT-based operational intelligence alert systems.



Object Detection




DaaS provides Object Detections of ground, air, and maritime military material, personnel, installations, and platforms, unit markers, banners, flags, and insignia. Methods for Object Detection are typically neural network-based or non-neural approaches. This platform utilizes a generic neural network-based CV model(s) to filter relevant content with a lower confidence. Neural techniques can perform end-to-end Object Detection without specifically defining features and are typically based on convolutional and recurrent neural networks.

Once a generic detection is accomplished, the platform employs an ensemble-supervised model trained to identify specific mission-related content relevant to the region and/or adversary. Confidence weightings are then applied to each relevant CV tag for cataloging purposes, resulting in highly accurate capture of mission-related imagery (with minimum false positives).

OCR Detection

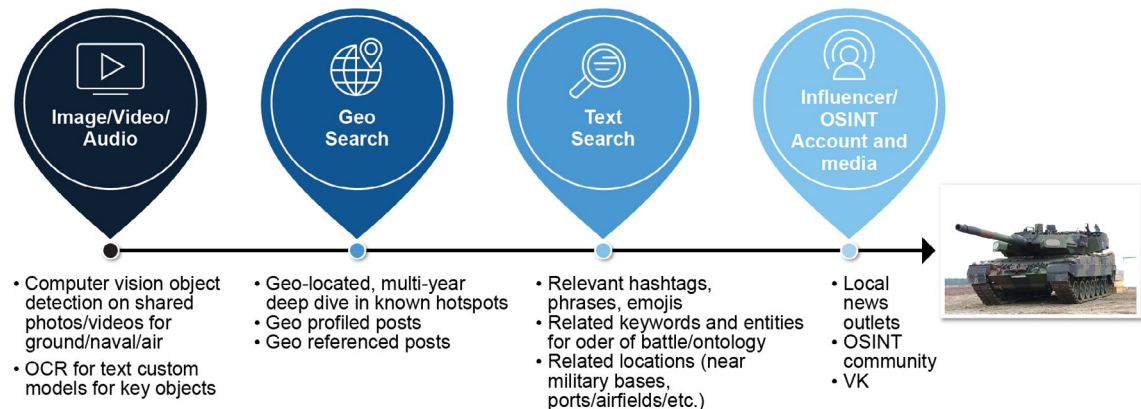
Optical Character Recognition (OCR) is the electronic or mechanical conversion of images typed, handwritten, or printed text into machine-coded text, whether from a scanned document, photo of a document, scene-photo, or from a subtitle text superimposed on an image. DaaS identifies and extracts nomenclature, equipment markings, license plates, and enable enlistment of these OCR detections as search filters and unique object tracking across the supported AOR(s) and geo-filtered based on Named Areas of Interest (NAI). The platform uses a standardized process to enable OCR detection, ensuring high quality results.



Impact & Benefits of DaaS	
	Scaling the ingestion and analysis of Russian and Chinese social media data of all AORs facilitates faster development and optimization of NRT alerts
	Application of AI/ML and proprietary algorithms limits human error and preset analysts with more useful, targeted information to produce intelligence and contextualize data from other data sources
	PAI and Social Media data is digested and analyzed in NRT. If PAI and/or social media data relevant to the mission is identified in a timely manner, critical operational intelligence can be discovered and actioned on quicker than traditional intelligence sources

NLP/Entity Detections

This platform manipulates and interprets natural language (speech and text) found in PAI using proprietary software to improve CV and Alert models. This platform employs its entity extraction models and various heuristics to detect, for example, individual people, personnel, organizations, places, equipment, and land, air, and sea vessels/vehicles. It provides NRT detections of mission relevant PAI as a service and, to garner greater value from this service, this platform provides ‘C3 Alerts,’ that combine high confidence imagery tags with contextualization, corroboration, and anomaly detection (i.e., spatiotemporal context), providing higher value to analysts than detections only.



Proprietary Differentiators

This platform the only open-source media intelligence platform that offers customers with real-time data feeds, Narrative Detection & Analysis, automated alerts via Peak Detection, Targeted Sentiment Analysis, and synthetic (or bot) detection via Automation Score. Below is a summary of our proprietary services and high-level overview of how, when combined, the DaaS services justify a Sole Source award.

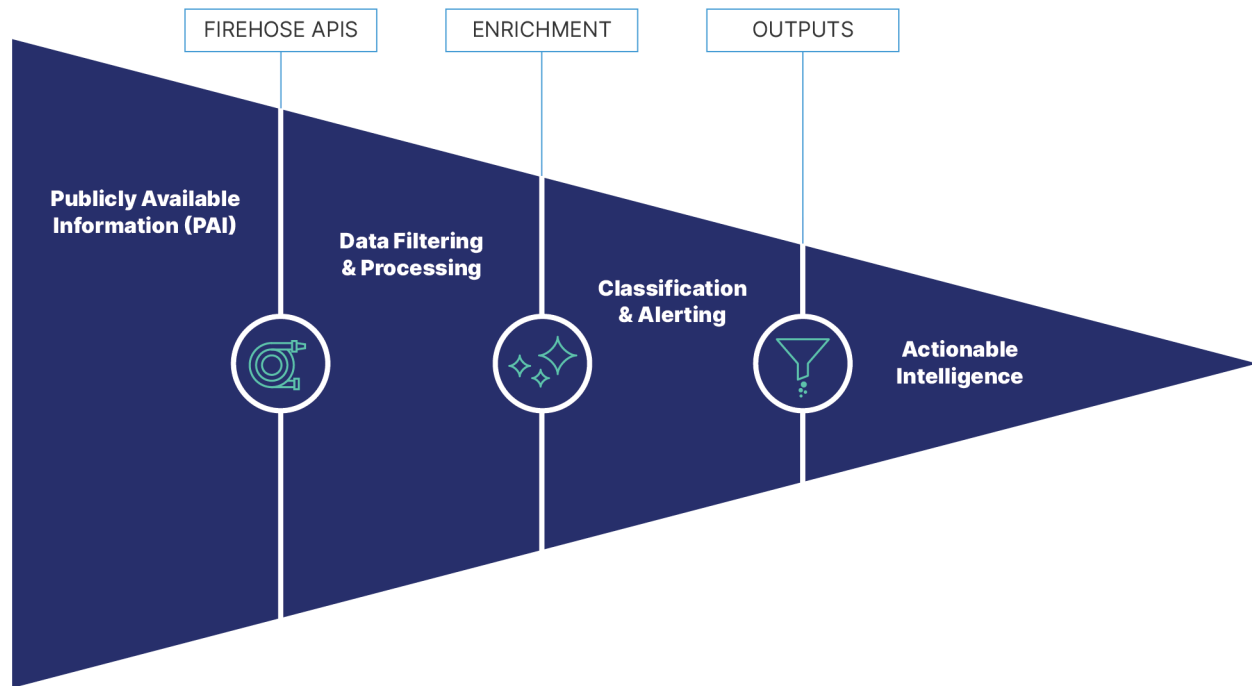
Data Accessibility

This is the only open-source media intelligence platform and digital analytics technology on the market that provides customers with data feeds, narrative analysis, and actionable intelligence in real-time on a single platform. Unlike other open-source data providers that merely scrape publicly available information, this platform constantly ingests billions of raw data and metadata points through direct data provider partnerships through full Firehose APIs.

Data Ingestion

The data ingestion model isolates and enriches open-source data through proprietary Artificial Intelligence and Machine Learning (AI/ML) technologies parsing through the noise and flood of available, allowing the platform to push only mission-relevant data to our customers. Our AI/ML technology equips our customers with visualizations and contextualizes the open-source digital media environment within a single software platform, providing a comprehensive view and specific narratives shaping global conversations, attitudes, and offline actions.





Augmented Data Discovery

This platform leverages neural networks to automatically extract and classify named topics and entities – people, places, products, organizations, or subject areas - from large datasets. Topics and Entities can be displayed in our OOTB visualizations, used as temporal search filters, and reapplied within overall search taxonomies. It also enables augmented insights through Restricted Boltzmann Machines (RBM) models that automatically surface and summarize the most representative stories in a conversation, allowing analysts to immediately understand the underlying narrative among thousands of documents

Automated Alerting

Manual searches and filters may not capture every piece of information an organization needs to act on emerging threats and opportunities. Thus, this platform pioneered proprietary AI/ML detection techniques for previously unknown or unpredictable changes in online conversations.



- **Peak Detection:** Through an understanding of historical data patterns, the DaaS automated early warning alerts proactively identify suspicious peaks in conversation virality, velocity, and volume.



- **Narrative Detection:** Unlike keyword-based searches and alerts, the platform's narrative-based detection model instantly surfaces media data points that share a common theme or agenda and form a particular 'narrative.'

Data Enrichment for Spot Analysis

Using Natural Language Processing (NLP) and rule-based heuristics, a select narrative or queried dataset is automatically enriched for geospatial inferencing, gender, influence, sentiment, automation, and more. The DaaS proprietary data enrichments allow analysts to conduct spot assessments of potential impact and develop parameters for subsequent investigation. These enrichments include:



- **Geospatial Inferencing:** In addition to ingesting user-reported geographic metadata across all media sources, the platform uses a proprietary model that leverages contextual information within a story or post to infer a location down to the city level. This model increases geographic coverage by at least 30% for sources such as Twitter, Instagram, traditional news, and broadcast television.



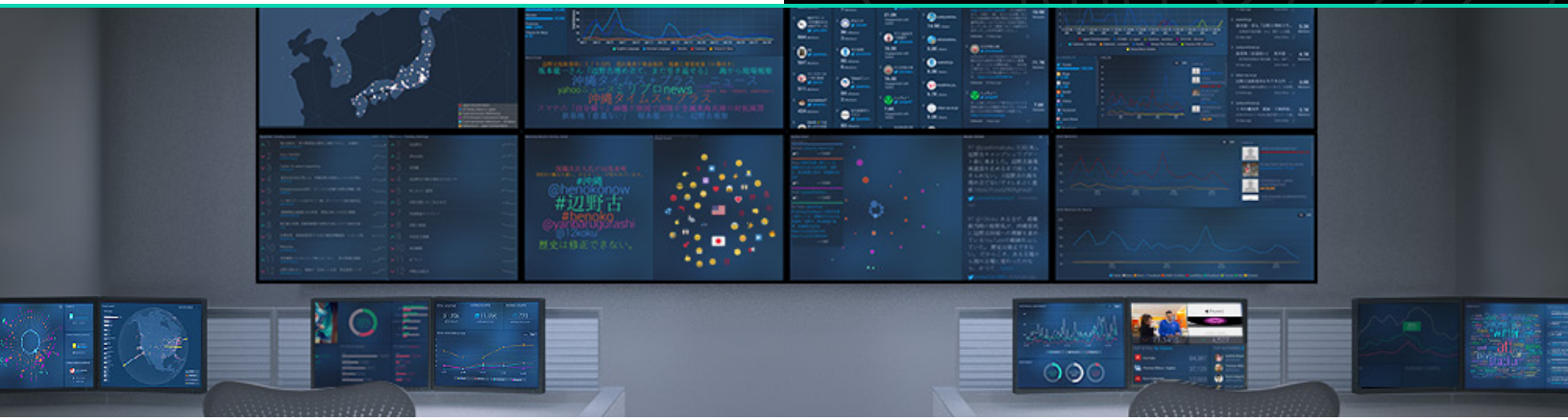
- **Targeted Sentiment Analysis:** The DaaS solution leverages a combination of proprietary classification system and a third-party partnership to provide sentiment scores in 78 languages and continuously evolve our models as the use of language changes across the world. The sentiment model uniquely targets 'mentions' of a specific organization, person, or issue within a story or post, and distinguishes the nuances of sentiment between mentions and that target and scores accordingly. While this patented model is highly accurate, the platform supports manual scoring override at the individual story level or in bulk for custom dataset tagging and classification.



- **Automation Score:** This platform was the first company to develop a proprietary automation score, which automatically categorizes authors based on heuristic behaviors, allowing users to quickly sort through the authors on a scale ranging from synthetic (or bot-like) augmentation to authentic and organic engagement. As synthetic propagation of content becomes more complex, the DaaS scoring scale provides a granular filter to understand a range of malign actor behaviors.



- **Influence:** Using metrics (i.e., author activity, follower engagement, and readership), this platform has created a proprietary scoring methodology to determine the influence of media outlets and social media actors. Within the platform, influence rankings can be visualized as an overall score or segmented individually based on engagement or impressions alone.



Dashboards & Visualizations

This DaaS has built numerous powerful and unique visualization capabilities, displayed via interactive dashboards, in real-time. Over 70 different visualizations – everything from word/emoji clouds to sentiment graphs and geo visualizations – can be included in a dashboard to align with a role, workflow, or challenge. All analytics visualizations can be explored in-depth by the end user. Data-drivers can click on any piece of information and drill down to the individual news story, blog, post, etc., to identify and expose the underlying story or social conversation. Visual analytics can be segmented by audience and/or influencer and organizable by topic and/or industry. Finally, Smart Filters allow users to apply filters with any combination of groups, issues, or metadata to the visual analytics for additional refinement and discovery.

Data Exporting

This platform allows users to export any/all collected data via TSV or CSV from either the dashboard or newsroom function. Exported data includes full text, metadata, and metrics, such as location, story subtype, bio, follower count, and more. For larger exports, it provides an API which allows for full text and metadata extraction as well as historical data. The API follows REST principles with JSON delivery to easily integrate with external systems. This platform allows for bi-directional integration via our programmatic administration endpoints.

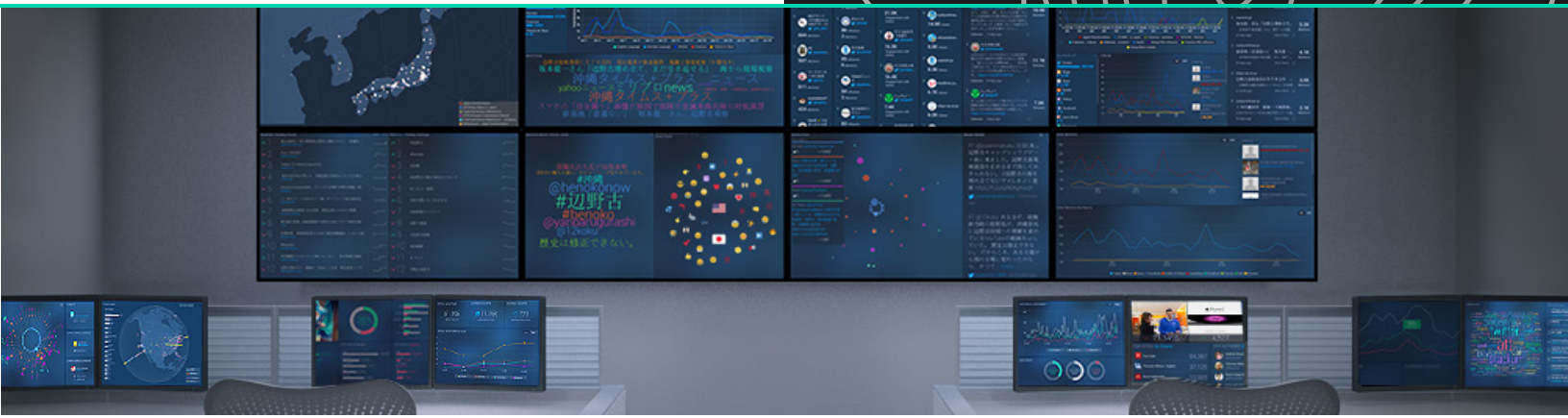
OSINT Global Pulse

Problem Statement

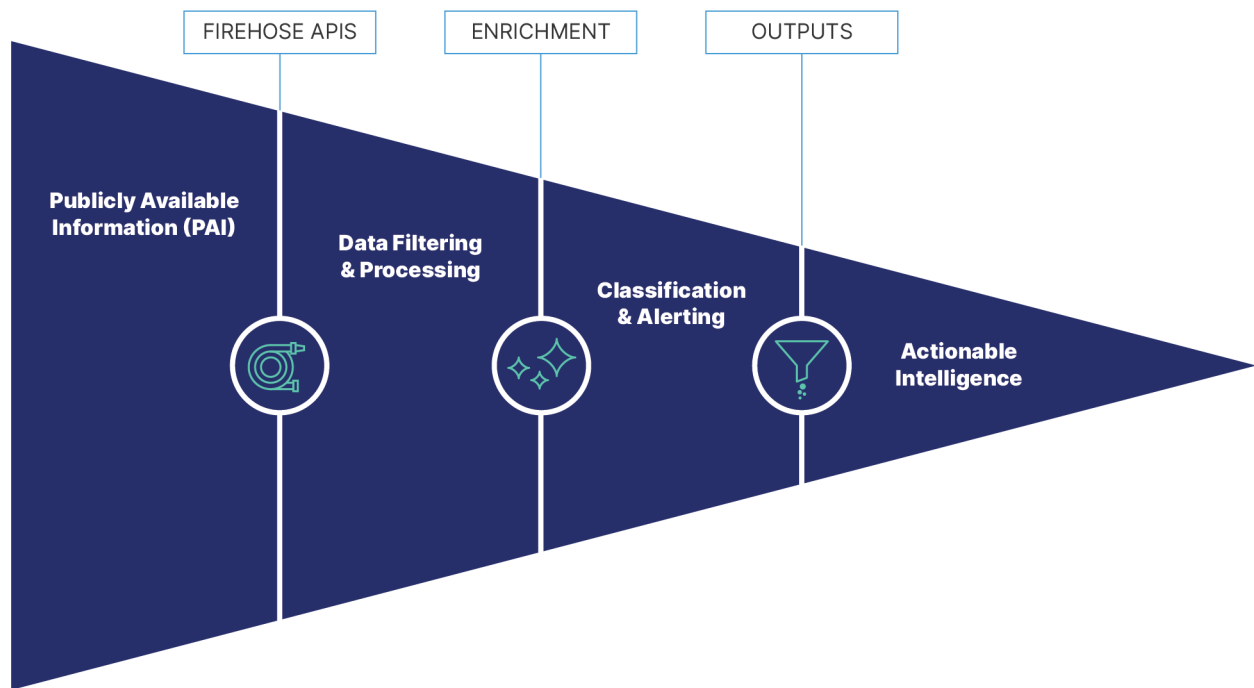
Digital and social media platforms are a new battleground for audience attention and perception. Adversaries are highly active in using these platforms to emblazon their narratives and encourage 'fires' to achieve chaotic effects. Conversely, neutral and friendly bystanders leverage global platforms to share their opinions, stories, and raise awareness on events impacting the safety of their communities. Although the rationale of both is in stark contrast, the output of each is vastly under-utilized; specifically, in the forms of aggregating sentiment, contextualizing images and videos, and internalizing their respective impacts on influence and intent. Equally critical to the understanding and analysis is the speed at which analysts, leaders, and decision-makers obtain mission critical perspectives of situational awareness. Operators in the information environment continue to struggle with obtaining definitive awareness in both the information and operation environments from the myriad of near-instantaneous information sources, particularly without an automated approach applying advanced AI/ML techniques.

Solution

This platform has critical experience transforming media intelligence into a strategic asset for civilian government, military, and intelligence organizations. With our original platform, ZEN, we implemented proprietary AI/ML technology designed to provide cutting-edge narrative intelligence, allowing analysts to detect individual media data points and prevailing narrative trends shaping the information environment in real-time. However, as global network infrastructure became available in regions once considered remote, and the technology supported on social media platforms evolved, the platform expanded its product offering to use the entirety of raw data and metadata elements posted by users. The newest ambitious offering, OSINT Global Pulse, is the industry's first product to expand AI/ML capabilities to include object, event detection, and deliver a contextual summary of events and associated data to end-users instantaneously.



OSINT Global Pulse is a next generation intelligence, surveillance and reconnaissance (ISR) tool leveraging publicly available information (PAI) Firehose APIs with many direct data provider relationships driving the largest social networking platforms in the world. This level of data access ensures nothing is missed in the information environment while preserving valuable metadata, allowing the platform to employ advanced analytics and proprietary AI/ML enrichment with industry leading outcomes. OSINT Global Pulse fuses Computer Vision (CV), Natural Language Processing (NLP), EXIF data, and narrative intelligence in Near Real-Time (NRT), providing a means to deliver critical alerts to emergency and force protection services derived from PAI.



Technical Approach

The DaaS solution ingests billions of raw data and metadata points from over 50 million unique global sources. These sources include, but are not limited to, full Firehose API access to Sina Weibo (SW), VKontakte (VK), and Twitter. These three social media platforms alone represent the largest social networks in China, Russia, and the wider world, respectively.















Standard Data Sources	Premium Data Sources
 News Online news articles from over 40,000 global outlets	 LexisNexis Global newspaper, magazine, trade journal, and other print articles typically behind paywalls
 Blogs Blog content from over 3 million global publications	 Twitter Public posts dating back to 2006 through full Firehose access
 Social Video Public videos posted on platforms like YouTube, Youku, and DailyMotion	 Facebook Public posts and comments on Facebook Page profiles
 Broadcast Full video clips and English and Spanish language closed caption transcripts from television stations in the United States, Canada, and the United Kingdom	 Instagram Public posts containing specific hashtags through a connected Business Profile
 Forums Public posts from global online forums and message boards, including alternative sites like Gab, 4chan, and 2chan	 Reddit Public posts and comments from the top 16,000 subreddit pages
 Q&A Full-text information from popular question-and-answer websites, such as Yahoo! Answers	 VKontakte (VK) Public posts on personal profiles and Community pages through full Firehose access
 Reviews User-generated review content from websites like Glassdoor and Amazon	 Sina Weibo Public posts through full Firehose access

FIGURE 1 - EXAMPLE OF DATA SOURCES

The ingestion model isolates and enriches data received with Exchangeable Image Files (EXIF data). The DaaS solution then uses proprietary AL/ML models to enable CV and NLP processing to assign pre-specified weighted labels and to detect narratives of concern, worldwide. Alerts are triggered in NRT using a rules-based heuristic scoring model, thereby informing analysts of suspicious activity, malign narratives, and geographic sources, while ultimately improving the Common Operating Picture (COP).

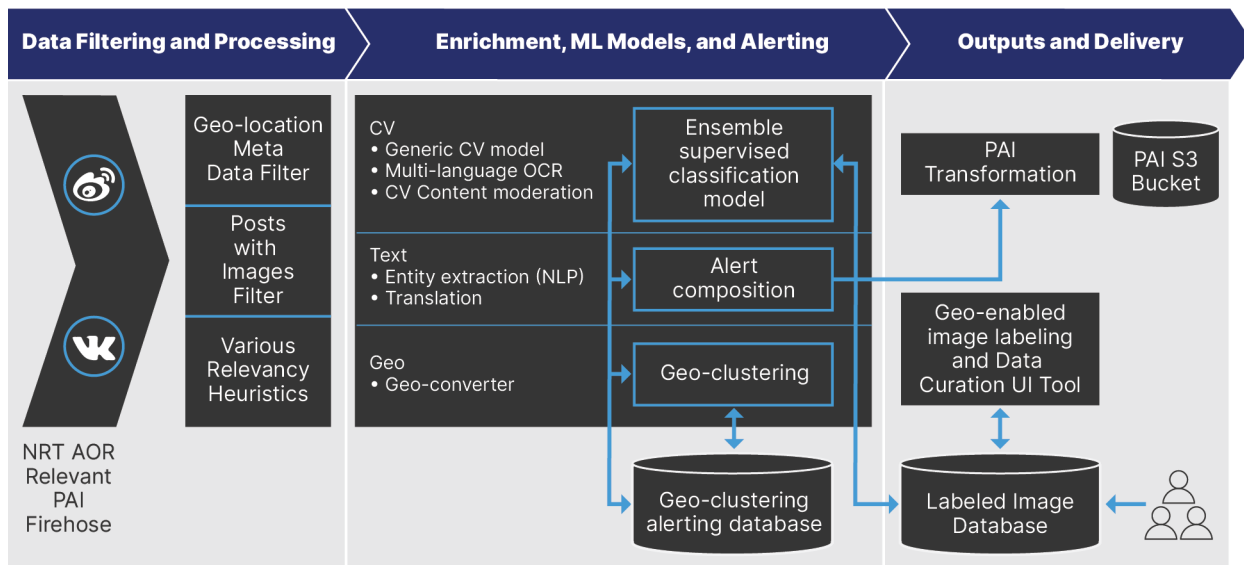


FIGURE 2 - DETECTIONS AS A SERVICE DATA INGESTION MODEL

Leveraging a combination of ML and image interpretation software, this solution identifies and isolates visual media (photos/videos) of relevance (removing false positives) to highlight candidates depicting, for example, military personnel, equipment, vehicles, and government/military installations. This platform also analyzes its recency and impact on the information environment through narrative intelligence analysis. Media impact is determined by examining environmental baselines and observing if there is heightened or abnormally high interest in imagery or a narrative. Measuring impact against a baseline provides a continuously evolving data set with immediate utility and applicability to analysts developing OSINT-based operational intelligence alert systems in NRT.

The enrichment algorithms are driven by priority intelligence requirements and continuously refined, increasing alert speed and accuracy rates, improving analysts' ability to understand, predict, respond, and/or influence decisions and associated behavior.

OSINT Global Pulse - EUCOM Examples

**** Imagery of Ukrainian bombing of Russian air base in Novofedorivka, Crimea ****

Imagery of Ukrainian bombing of Russian air base in Novofedorivka, Crimea

2022-08
45.0955278, 33.5763273 Imagery of Ukrainian bombing of Russian air base in Novofedorivka, Crimea

2022-08-09 14:16:25 GMT
Big explosion in Crimea leading to a fire. Former advisor to the head of MIA of Ukraine saying it's a long-range missile, but unconfirmed for now
POV: 45.127941, 33.520457

Center: lng 34.7937 lat 45.7832
Radius: 15.30345 km
© 2022 Confidential

U.S. GOVERNMENT DISTRIBUTION ONLY

EUCOM EXAMPLE 1 - UKRAINIAN BOMBING OF RUSSIAN AIR BASE

**** Wagner position in Popasna struck by HIMARS ****

A Russian Telegram channel confirms that there was a strike on a Wagner position in Popasna likely by HIMARS. The position of the Wagner soldiers was leaked to Ukrainian military by a Telegram user account that shared the photos of Prigozhin in Popasna, and later deleted the posts. There is also a video showing the arrival at the location of Wagner PMC in Popasna from taken by a person who states 'Slava Ukraine' while taking the video hiding nearby.

2022-08-14 16:42:08 GMT
@NTenzel The Armed Forces of Ukraine bombed the headquarters of the "Wagner Group" in the occupied Popasna Luhansk region.
<https://t.co/Zw9u7HZV5a>

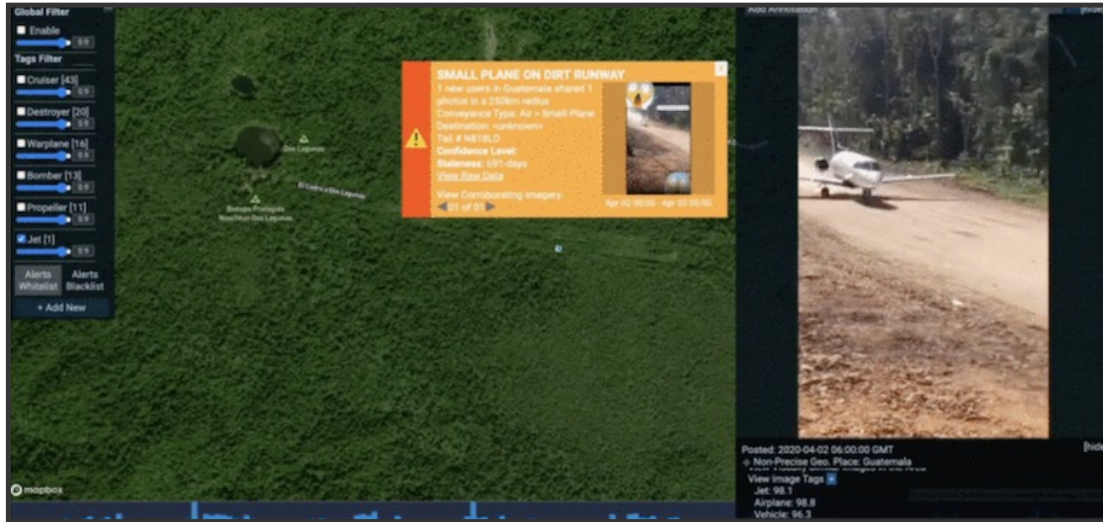
2022-08-14 17:52:47 GMT
Video after arrival at the location of Wagner PMC in Popasna <https://t.co/s9YL8ovt4>

Center: lng 38.3812, 48.6332
Radius: 6.99 km
© 2022 Confidential

U.S. GOVERNMENT DISTRIBUTION ONLY

EUCOM EXAMPLE 2 - RUSSIAN CONFIRMATION OF WAGNER POSITION

OSINT Global Pulse - SOCOM Examples



SOCOM EXAMPLE 1 - NARCO AIRCRAFT IN GUATEMALA



SOCOM EXAMPLE 2 - SUMMARY OF AIR & MARITIME NARCO EVENTS

<p>Title: Mexican Military Captures Narcoplane in Campeche, Mexico Body: The Mexican military captured a narcoplane with approximately 460 kilos of cocaine in La Libertad, Campeche, Mexico. Created At: 2022-08-15T22:04:46Z Range: 2022-08-14T18:36:58Z, to: 2022-08-14T19:16:47Z Proofs: post: 3, image: 7, Corroboration: post: 0, image: 0</p>	<p>Title: Colombian and Ecuadorian Authorities Captured Narcoboat near Esmeraldas, Ecuador Body: Colombian and Ecuadorian authorities captured a narcoboat with 378 drug packages and arrested three individuals near Esmeraldas, Ecuador. Created At: 2022-08-15T21:56:44Z Range: 2022-08-13T20:10:28Z, to: 2022-08-14T13:20:47Z Proofs: post: 2, image: 2 Corroboration: post: 0, image: 0</p>	<p>Title: Oil Terminal Fire in Matanzas, Cuba Continues Body: Oil Terminal Fire in Matanzas, Cuba the night of 5 August 2022 results in 2,000 evacuated and over 120 injured. International aid incoming from Mexico and Venezuela to combat the spread of the fire. Created At: 2022-08-09T15:16:33Z Range: 2022-08-06T00:39:01Z- 2022-08-06T16:52:20Z Proofs: post: 5 image:7, Corroboration: post: 0, image: 0</p>

SOCOM EXAMPLE 3 - AIR & MARITIME NARCO EVENTS

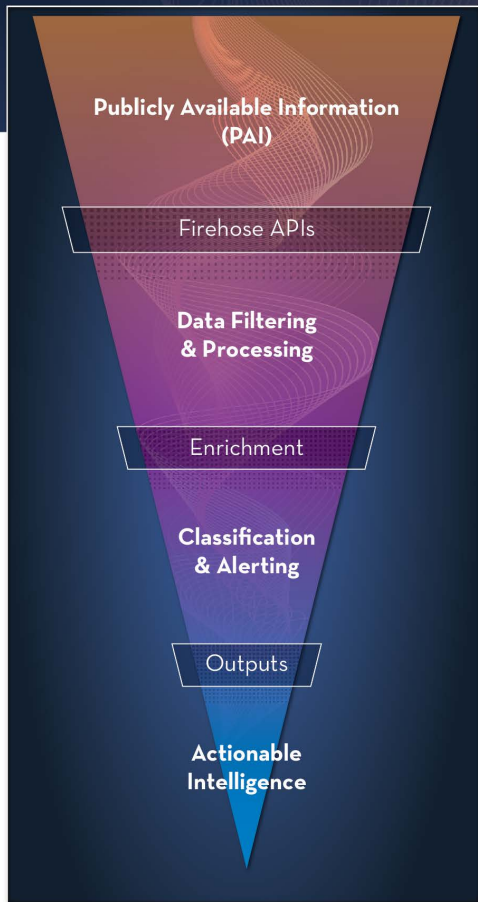
Conclusion

Social media has revolutionized the means and modes of communication, from the platform a person, group, or global leader chooses to use, to the matter in which they do so; constantly evolving the use of language, context, and intent, not only through narratives, but through pictures, emojis, videos, and in every combination thereof. Although OSINT can provide global insights, the current methods of data collection, analysis, and production are unscalable and severely handicapped when it comes to continued reliance on human analysts and the availability of timely data.



Our Mission

To implement Artificial Intelligence (AI) and Machine Learning (ML) solutions, enabling Near Real-Time (NRT) delivery of Computer Vision (CV) and Natural Language Processing (NLP) object and narrative Detections as a Service (DaaS) via Publicly Available Information from global media platforms.



Our Strengths

Comprehensive View

- Visualizes & contextualizes the digital media environment
- Easily detect specific narratives shaping global conversations

Media Monitoring

- Establishes patterns & techniques used to manipulate discussions

Emerging Narratives

- Early warning alerts
- Identify peaks in conversation virality, velocity, and volume

Full Firehose APIs

- Direct data provider relationships
- Bi-directional API's
- 50+ unique sources including Sina Weibo, VKontakte, and Twitter

Data Ingestion Model

- Ingests billions of raw and metadata points
- AI/ML models enable CV & NLP
- Integrates Narrative Intelligence
- Precipitates I&W Alerts

Actionable Intelligence

- Push enriched metadata via REST API
- Available in near Real-Time



Peak Detection



Emerging Narratives



Geospatial Inferencing

Proprietary Technology



Sentiment Analysis



Automation Score



Influence Intelligence

U.S. Defense Customers



U.S. Federal Civil Customers



Commercial Clients

