CLARITAS

REDEFINING INTELLIGENCE

BLUEFUSION is a mission focussed Intelligence platform designed to deliver rapid, actionable strategic and tactical insights for intelligence and supporting missions.

BLUEFUSION uniquely provides a unified, semi-automated, truly multi-data and cross domain capability with a global coverage, for continuous person-centric and high-volume multilingual thematic collection across the surface, deep and dark web, social media, forums, emerging platforms and data-broker sources with entity ingestion, link analysis, scalable keyword campaigns, real-time alerting, and discovery of emerging trends and Organised Criminal Group (OCG) modus operandi.

We cover the deep, dark net alongside market places, access brokers, chat forums, messaging apps and others, our truly unified OSINT platform is purpose built to to enrich the subjects of interest and discover additional intelligence, as well as automatically ingesting and matching data such as Call Data records against SOI where this data is available.

USP

Leveraging advanced AI/ML alongside multi-source data fusion, our unified flagship platform seamlessly brings together proprietary, open-source and commercial datasets at scale and pace. No competitors can connect visual, device, telecom, digital, and movement data into one cohesive system and at this scale. Few providers can uniquely deliver an integrated, end-to-end operational workflow — from physical collection through to AI-generated insight and action fuelled by a truly global passive collection of commercially available data.

Our core differentiator is the combination of unique capabilities that competitors can only deliver in isolation:

Deterministic identity resolution — ADID ↔ MSISDN matching
Hash cracking at scale
Massive, queryable data lake — no API limits, no sampling, no external dependency
AI operating across the full owned dataset — not cloud-dependent, not sampled, no partial sets
Sovereign, in-country deployment if required — compliant with the strictest data residency requirements
Fixed cost, predictable, annual licensing.

This is extremely difficult to replicate — it requires proprietary IP, data access, infrastructure, deep integration, and mission-domain expertise accumulated over years.

Data Scale & Coverage

we do not offer up sample datasets, nor do we offer partial datasets, nor do we offer up API callouts or limited queries.

We provide entire, complete and pre-processed AdTech datasets that are also fused with complimentary datasets to maximise the enriched intelligence value possible alongside providing a high degree of confidence for users and decision makers for over 200 countries.

We ingest Trillions individual data points, from multiple suppliers all over the world and is subjected to a rigorous and highly tuned forensic flagging process.

For unique device SDK data, in April 2026 we visualised 2,787,614,522 cleaned, normalised and quality checked unique and verified device data points as monthly averages.

With regards to unique ADID Signals, in April 2026 we visualised 482,839,792,519 cleaned, normalised and quality checked unique device data points.

Our Hashed Email (HEM) records in April 2026 we visualised 634,919,168 cleaned, normalised and quality checked unique data points as monthly averages.

Our Device records in April 2026 we visualised 31,369,735 cleaned, normalised and quality checked unique device data points as monthly averages.

Our App Unique records in April 2026 we visualised 485,159,594 cleaned, normalised and quality checked unique App data points as monthly averages.

Our App Device records in April 2026 we visualised 84,142,204 cleaned, normalised and quality checked unique device data points as monthly averages.

This excludes the API data being ingested for AIS, ADSB and other unique feeds we are ingesting daily, for example OSINT feeds. We are also able to run commercially provided RTB campaigns as part of any low latency requirement in support of urgent or emergency operational requirements.

This translates on average to around 300TB of hot data storage and cataloguing we are providing at any one time.

Analysis

We provide a workflow to search global historical SDK data as recent as 2-3 days old and as far back as 2 years old. Users may pull millions of events for fine-grained manipulation and analysis, including the ability to replay device movement in granular detail, in chronological order.

This allows identification of patterns, behaviours, co-travellers and trends over any given period of time, adjusted as data is discovered. AI also identifies indicated behaviours automatically, for example the device user is most likely driving.

Sourcing & Quality

SDK data delivery is delayed primarily due to sourcing and aggregation due to processing and transfer. SDK data in the data lake is lawfully sourced from over 7 different vendor sources, which are all de-duplicated, analysed and flagged for different quality metrics in over 200 countries. Users are able to apply filters on these quality metrics to decide when to include them in analyses or to exclude them.

Individual signal records that are within 2-meters in space and 35-seconds in time are consolidated before delivery, sourced from multiple applications from a variety of publishers, all from different points of origin.

The average consolidation/merge factor for all data is ~1.58. Therefore, if the SDK data for a country has 100M lines, then the record count is ~158M individual signals prior to merging.

This approach to signal consolidation saves customers from ingesting, processing, storing, and querying 7 quadrillion rows of data annually, making us highly efficient and cost effective with the ability to provide a fixed annual licence fee for the core system implementation and delivery.

Location metadata for IP intelligence also runs at 100% success rates and is scored by confidence in accuracy from a national to a local level and the platform supports strategic reporting, trend analysis, dashboards, reporting outputs and broader assessment activity as standard.

Forensic Flagging

Location data sources include GPS, Wi-Fi, cellular, IP-derived and other relevant signals and accuracy is assessed and confidence scored based on forensic flagging. Statistically, our flagging, cleaning and normalisation processes when analysed against the above ingestion figures sees exceptional success rates.

Quality flags on our AD-ID SDK fused Data include:

- Radio Derived Location. Detects and flags locations that were likely generated from Cell Tower locations, not precise GPS location of handsets.
- IP Derived Location. Detects and flags locations that were likely generated from looking up the IP address associated with the signal, rather than the handset GPS..
- Too Many Devices at a Location. Flags records involved in too many different devices reporting the same latitude/longitude in a day. This may not necessarily be indicative of bad data or bad activities, but it is flagged as something interesting to investigate.
- High Activity. Detects and flags devices that emit lots of SDK location data in a short period of time.
- Grid-like Location. Flags locations that appear to fall in a grid pattern for a device, which is indicative of estimation techniques from app developers.
- Spoofed Locations. Flags locations that appear to have been synthesised based on a generic coordinate used by an app, or one that is made up, like (0.000,0.000).
- Improbable Speed. Flags records that report locations that are too far apart, in too short a time span to be real.

Recurring Patterns. Flags records that appear to be simulated via playback of the same movement pattern and speed.

ADID Search and Correlation

Search by Ad-ID to discover associated other identifiers used by the device, such as email addresses and MSISDN, search using postcodes and addresses and filter devices based on frequency of visits to locations/ perimeters. Automatically associated them with SOI within the system, populating the visualised investigation automatically with any other related data, such as call data records, co-travellers, HEM etc, that is deterministically shown to be associated with that specific selector - without any input required from the user. Hashed values can also reveal clear text email from certain datasets, including unhashed ADID and HEM we provide and continually update as standard at no additional cost.

Co-Traveller AI

Global search enables users to access co-traveling devices, ships and planes that the system has automatically discovered among all its input data feeds. With AI running operating across our own datasets, means no reliance on cloud based AI that can fluctuate during periods of high demand, delivering unique datasets and correlations across multiple locations, able to be exported in standard formats for ingestion into other systems.

The system will automatically Identify previously unknown devices or associations through correlation of advertising IDs across datasets and locations to populate subject of interest co-travellers automatically.

Interface

The Analysis tooling is designed with user friendly Analyst centric interfaces. The map-centric investigative view provides another way for users to search for Ad-ID data based on defining geographic places and polygons on a map, and even to find common Ad-IDs found intersecting multiple geographical areas via Geofencing and time ranges as well as inclusion of whole country blocks, groups or bespoke lists of specific data and categories to allow Analysts complete flexibility as to how they discover the data.

Result filters are also available to further refine searches based on other fields such as horizontal accuracy, carriers, device models, and more which can also be dynamically refined quickly and easily, for example by a specific type, category, time, pattern of life, carrier etc to reduce noise quickly and effectively.

Accuracy

The datasets on offer are refined and granular, enabling Analysts to drill right down to isolate devices of interest. Results can be prioritised and precision may be as accurate as 1-10m with variances in the hundreds of meters pending many factors. For example, devices without clear skyward line of sight will typically not achieve 1-10m geo-data resolution. Additionally, various geo-resolution methods can approximate location based on neighbouring signals such cell-tower location and IP address clustering. We use forensic flags embedded in each signal to provide insights into the derivation and quality of the signal including High Accuracy or Approximate location and ensure this assessment is always visible to the analyst.

Data accuracy information is embedded into analytical outputs to support confidence in findings, for example a % confidence in an IP address geolocation accuracy, or a match with a selector where applicable, with the category and type visible at all stages so the source is visible i.e in the case of a satellite data source it would be attributed specifically to Iridium.

With regards to the statistics on the compliance and performance monitoring, we have no doubt that BLUEFUSION will have no problem achieving the stipulated thresholds.