ENEA

SOLUTION BRIEF

Qosmos for Cybersecurity:
Rich Network Traffic Data for Machine Learning

Because the Success of Your Solutions Depends on the Quality of Your Data

Key Benefits

Quality Data for Better ML & AI Results

Qosmos is committed to delivering real-time traffic data of unmatched depth, breadth and accuracy. The data is clean, structured (feature-value pairs), and ready for use in supervised and unsupervised machine learning projects.

- **Differentiate Your Offer**
  Use Qosmos' exceptionally rich and comprehensive traffic data to produce ML results that distinguish your offer through greater accuracy and relevance.
- **Boost Analyst Productivity**
  Benefit from Qosmos' built-in data mining tools and processors to dramatically reduce the huge amount of time your data scientists spend gathering and preparing data.[3]
- **Increase Stakeholder Confidence**
  Leverage Qosmos' reputation for top quality traffic data to strengthen stakeholder confidence in your solutions.
- **Support Continuous Innovation**
  Take advantage of a broad, diverse repository of traffic data to address new use cases and emerging threats.
- **Benefit from Real-Time & Historic Data**
  Take advantage of Qosmos' real-time data at scale to continuously refine your model, and use it to build a resource-efficient archive for new model development.
- **Detect Threats Earlier & More Accurately**
  Use Qosmos to help deliver ML and AI solutions that make the most of the innate data advantage your customers have, and deliver the best possible protection against advanced persistent threats.

CYBER SECURITY EXCELLENCE AWARDS ★ WINNER ★ 2019

While model efficacy (and the data quality upon which it depends) is important in every industry, it is perhaps uniquely so in cybersecurity. Why? Because cybersecurity is locked in an arms race with sophisticated adversaries that are themselves using machine learning (ML) and artificial intelligence (AI) to thwart defensive systems. The financial, and increasingly physical and human, costs of failure are extraordinarily high.

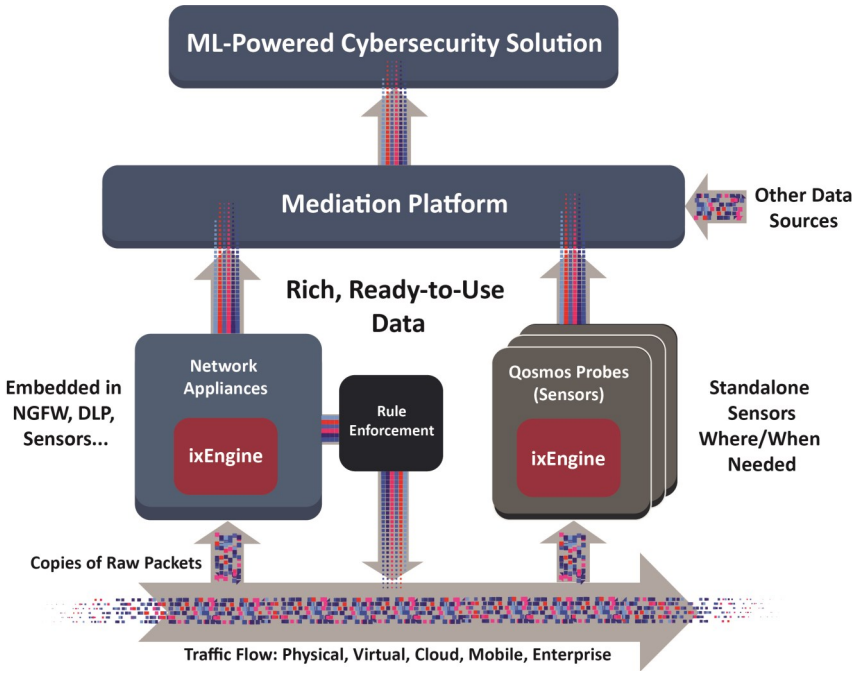
In a recent Dimensional Research survey[1] of data scientists and other AI professionals in large companies across 20 industries worldwide, nearly eight out of 10 organizations reported that their projects had stalled or been aborted. An overwhelming 96% cited challenges with data quality, the data labeling necessary to train AI, and building model confidence.

These findings may seem astounding, but are not surprising to ML practitioners, nor to researchers. The latter have demonstrated conclusively that data quality has an enormous effect on the accuracy and efficacy of ML results, so much so that enhancing the quality of the input data alone can dramatically improve a model's output—without making any changes to the algorithm (or algorithms) used.[2]

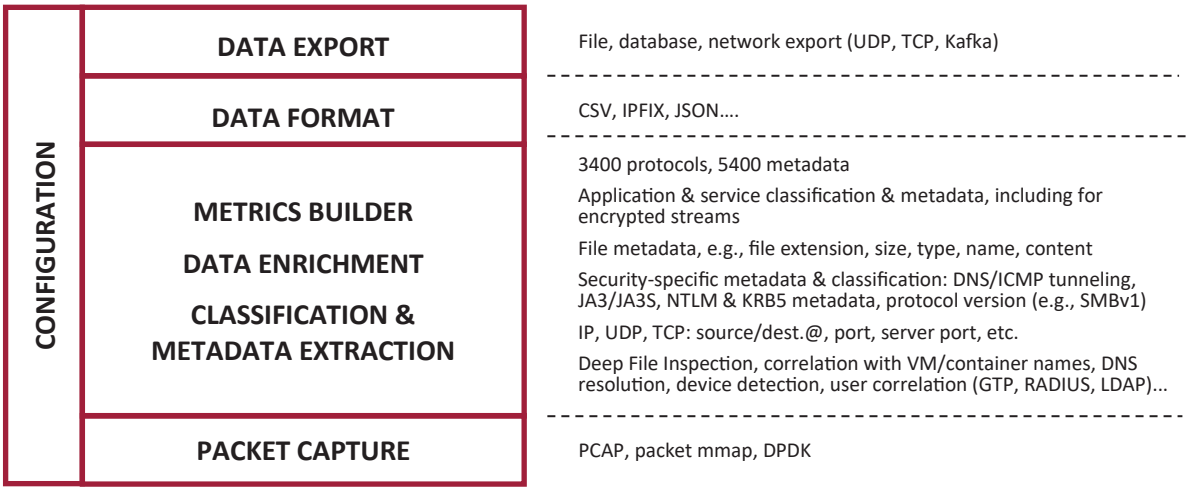
So, what can be done?

Cybersecurity has a huge advantage in the data it can access. While the good guys and the bad guys both have algorithms and experienced analysts and programmers, the good guys also have access to an abundance of unique data on end devices, users, equipment, applications, infrastructure and traffic flows that attackers do not (short of a mass exfiltration of all data from all systems).

Exploiting this advantage only requires making maximum use of the volume and variety of this data, and ensuring the quality of data from each source. Network traffic data is arguably the most important input source for cybersecurity (especially when correlated with endpoint logs), and it is here that the Qosmos ixEngine® is your most powerful asset.



Qosmos ixEngine & Probe Functional Architecture



About Qosmos Technology

The Qosmos ixEngine uses deep packet inspection (DPI) technology coupled with flow analytics and advanced data mining techniques to deliver real-time traffic data of unmatched depth, breadth and accuracy. And, the data it produces is clean and ready to use in supervised and unsupervised machine learning projects.

Below are examples of the analytic techniques Qosmos uses, most of which are executed in tandem to produce the most accurate results.

- Handshake Analysis
- Traffic Pattern Analysis
- Statistical Modeling
- Behavioral Analytics
- Predictive Session Analytics
- Session Correlation
- Public IP/Port-Based Correlation
- Deep File Inspection

About the Qosmos ixEngine®

Available as an SDK (in C) for embedded use, and as a standalone application (the Qosmos Probe), the Qosmos ixEngine is the most widely deployed commercial traffic data engine in cybersecurity and telecommunications. It excels at transforming raw packet and flow data into uniquely enriched (and customizable) data streams for operational and analytical security components and platforms.

Data Quality

Qosmos ixEngine is the industry gold standard for data quality. The traffic data it produces is:

- **Accurate**
  It is based on the most trustworthy source available - telemetry data (not insecure log files), and rigorously validated. It is this reliability that enables Tier 1 vendors to use it with confidence in their planning, decision-making and operations.
- **Comprehensive**
  It includes data on *all* network flows provided, and produces highly granular data for these flows. In total, the Qosmos ixEngine provides classification data for 3400 protocols (including IoT/SCADA & cryptocurrencies), and produces 5400 different types of packet and flow metadata—the broadest coverage in the industry
- **Relevant**
  It is precise, contextual data delivered via a framework that offers maximum flexibility in selecting the data features most relevant to the goals of your analysis, and in defining rules based on that analysis
- **Real-time**
  It is generated from raw data captured on-the-fly via passive physical or virtual network TAPs that do not affect traffic flow
- **Always Up to Date**
  Updates are continuous and hot-swappable to ensure you will always stay abreast of constantly changing applications and protocols, and benefit from the latest advancements in data classification, especially for encrypted and evasive traffic

Types of Metadata Produced

Beyond IP traffic classification, the Qosmos ixEngine® extracts 8 main categories of network-based application and computed metadata:

- **Volume**: e.g., the volume of traffic per application and per user
- **Service identification:** e.g., service classification for VoIP and IM protocols and applications, even in encrypted streams
- **Application usage:** e.g., SMB:version, user_agent length (for entropy), file hash
- **Identifiers:** e.g., email sender / receiver addresses or any other ID that can be used to implement strong security rules
- **Content:** e.g., link detection and extraction in email; attached files in email, which can be directed to specific processing like 3rd party anti-virus or content inspection
- **File metadata:** such as file extension, size, type, name, content, etc., which can be very helpful for use cases such as DLP or advanced malware detection
- **Security-related classification & metadata:** e.g., tunneling on protocols such as DNS or ICMP, NTLM & KRB5-related metadata, JA3/JA3S hash to fingerprint TLS connections, protocol version (e.g., SMBv1)

Learn More

For further information, visit the Products section of our website at: https://www.qosmos.com/products/

(1) https://bwnews.pr/2UPcZZE
(2) See for example, https://bit.ly/3bzgyJ9
(3) Data scientists spend up to 80% of their time gathering & preparing data: https://bit.ly/2ORuSDf

ENEA

Enea is a world-leading supplier of innovative software components for telecommunications and cybersecurity. Focus areas are cloud-native, 5G-ready products for mobile core, network virtualization, and traffic intelligence. More than 3 billion people rely on Enea technologies in their daily lives.
For more information on Enea's Qosmos ixEngine, Qosmos Probe or Qosmos DPI technology: www.qosmos.com

Find out more!

www.enea.com