QOMPLX:OS

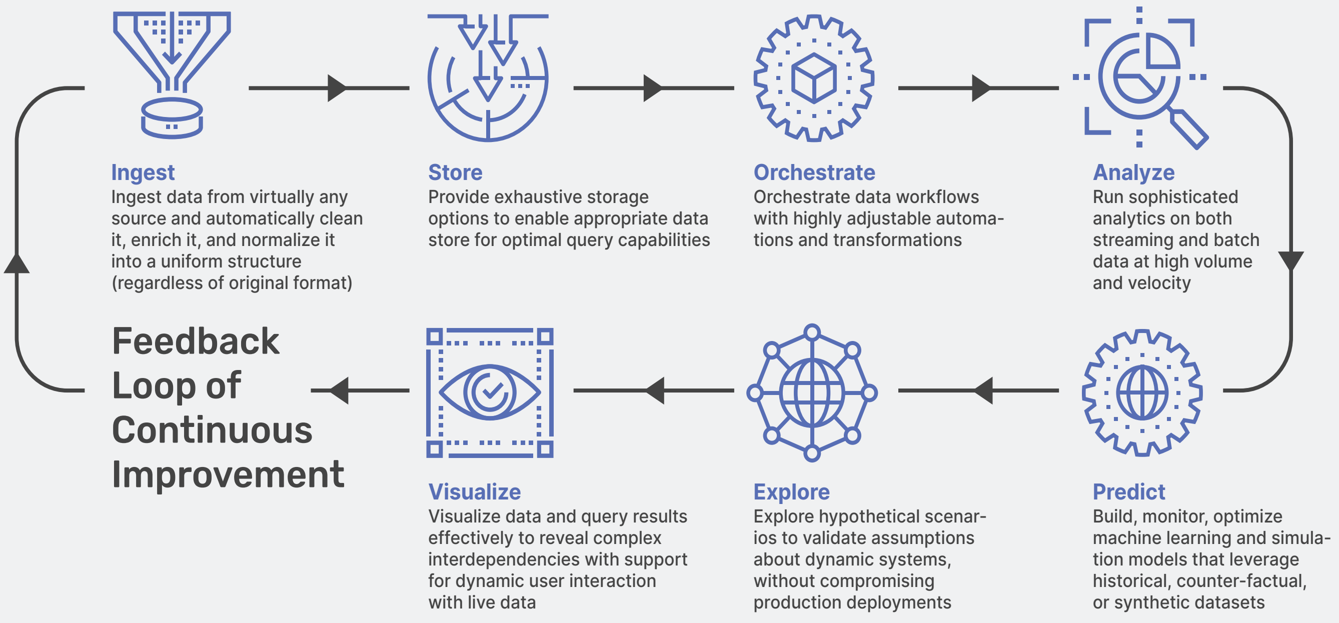
Leveraging a Decision Platform for Meaningful Analytics

Why QOMPLX?

The world is moving forward at lightning speed—more and more interconnected devices and systems are finding their way into our lives and powering our work. This fundamental shift is resulting in exponential growth in data production, with increasingly different data types from vastly different sources. Our ability to reason effectively with this deluge of information to make thoughtful decisions is directly related to our ability to properly consume such massive amounts of data.

Fortunately, QOMPLX has created purpose-built technology in QOMPLX:OS (Q:OS) to ingest, analyze, enrich, and report on data to enable contextualized decision-making at any scale. While particularly focused at the nexus of cyber, financial, and operational data, the core analytic building blocks and highly customizable workflows available within the Q:OS platform can be recombined to support data-driven decision making, in any domain. It all starts with QOMPLX's entirely innovative approach which prioritizes up-front effort to ingest, schematize, normalize, and semantify data in order to impart context to any data source, regardless of its original format.

An End-to-End Data Analytics Platform



Ingest
Ingest data from virtually any source and automatically clean it, enrich it, and normalize it into a uniform structure (regardless of original format)

Store
Provide exhaustive storage options to enable appropriate data store for optimal query capabilities

Orchestrate
Orchestrate data workflows with highly adjustable automations and transformations

Analyze
Run sophisticated analytics on both streaming and batch data at high volume and velocity

Predict
Build, monitor, optimize machine learning and simulation models that leverage historical, counter-factual, or synthetic datasets

Explore
Explore hypothetical scenarios to validate assumptions about dynamic systems, without compromising production deployments

Visualize
Visualize data and query results effectively to reveal complex interdependencies with support for dynamic user interaction with live data

Feedback Loop of Continuous Improvement

Decomposing Data for Analytics

The real world requires that teams can easily begin ingesting their data from a plurality of sources, adapting to the rate and volume at which it is being created. To properly prepare the data for analysis, it is typically deduplicated to reduce dataset size and streamline it for future processing. This often involves leveraging various mathematical formalisms and advanced natural language processing (NLP) algorithms for unstructured data to recognize and remove redundant information. At this point the data is stored in a model optimized for querying and analysis, such as a graph, time series, index, relation, or key-value store. For example, a block of text will be broken up and placed in rows in a table for faster processing but a highly relational data set may need to be in a graph to support efficient traversals.

Once the data is collected, deduplicated, and properly arranged, it is normalized by converting various data types into a consistent unified data model. It is important that data is correctly associated with appropriate entities and that units or values are uniform, perhaps converting all temperatures to either Celsius or Fahrenheit or all time to the GMT time zone. This process further prepares the data for analysis with various predictive models, including machine learning algorithms. Depending on the use case, it can be performed either on streaming data as it is ingested or on batches of stored data.

Next, a standardized set of naming conventions or ontologies are applied to the data to ensure there is no confusion about the definitions of objects or values within the dataset. The process of indexing data points with their logical or linguistic meaning in the context of surrounding data points is referred to as *semantifying* data. It facilitates deeper reasoning and automated analysis of entities within the data and their relationships to one another.

Automated Data Management for Streamlined Processing and Analytics

Ingest

Next-generation data extraction capabilities allow the ingestion of more structured and unstructured data from disparate sources in batches of streaming

Deduplicate

NLP algorithms can be leveraged to recognize and correct duplicate and/or redundant data

Schematize

NLP algorithms can be leveraged to recognize and correct duplicate and/or redundant data

Normalize

Real-time data processing capabilities allow data to be rendered uniformaly (e.g. real-time conversion of all currencies to US dollars)

Semanify

Standardized ontology enable machines and people to understand, share, and reason with information efficiently at execution time

Modeling Data to Extract Meaningful Insight

Sometimes simple expressions and logic are enough to glean sufficient insight from this semantified data, but other times more advanced techniques involving various machine learning or deep learning approaches may be needed. For example, model-driven analysis extracts intended meaning and sentiment from widely varying sources of structured or unstructured data, allowing infusion or enrichment of the semantified data with even clearer meaning and greater context for increasingly improved automated processing and more insightful interpretation.

Data can now be leveraged to build dynamic models that enable better understanding, prediction, or simulation of the world. Additional self-improvement for models may leverage orchestration tools or reinforcement learning techniques to recognize and correct for model bias to continually maximize the accuracy and effectiveness of those models over time. Contextualized information can now be presented in a way that enables Human+AI collaboration to optimize decision-making and performance, at scale, to any domain. Using the right model at the right time and in the right place is a priority.

Contextualized information can now be presented in a way that enables Human+AI collaboration to optimize decision-making and performance, at scale, to any domain.

Legitimate Integration of Disparate Security Data

Many of the techniques used by QOMPLX are domain-agnostic, meaning they can be applied to virtually any structured or unstructured data set. This is largely facilitated by Q:OS's ability to restructure these diverse data sets using Domain-Specific Languages (DSLs) and common data formalisms. In cyber, emerging industry standards such as Open Graph of IT (OGIT) are extended, allowing users to define enterprise-grade topologies of corporate networks; Structured Threat Information eXpression (STIX 2), enabling ingest and search of the machine readable threat intelligence reports; and even the Financial Industry Business Ontology (FIBO), facilitating more detailed exploration of how companies, people and other risk factors are connected. For most organizations today, security data and cloud telemetry, due to its abundance and pervasiveness, often act as the fountainhead for gaining the insight needed to understand a number of operational risks across the organization. Visibility into networks, and ultimately business processes, is critical to keep the CISO and the security organization connected to the business.

The effective collection of security data provides the foundation for understanding risks surrounding users, devices, and services throughout the enterprise. It is common for security teams to lack sufficient instrumentation and collection of data from a sufficient number of sources to accurately contextualize data. A growing set of tools naively advocate for use of a single data source (e.g. interior switch on an external gateway) as sufficient for advanced analytics and machine learning. However, the overstatement of capabilities from single-source analytic approaches are a major contributor to alert fatigue and noise in the Security Operations Center (SOC). Such tools simply fail to integrate and sufficiently correlate diverse data with proper context to support sound decision-making.

Extending Security Data Models to the Business

By organizing data into a unified model upon ingestion, Q:OS enables databases to evolve into knowledge bases where querying and analysis is intuitive and efficient, even when reasoning about data from heterogeneous and disparate sources. Actual insights are delivered across data sources and not just a storage bill associated with a large but often unusable data lake. With its cloud-distributed, highly parallelized analytic routines, Q:OS is able to process logs and instrumented interactions at massive scale and in near real-time to provide the immediate context needed to understand what's happening on the network when it's happening. As a result, its process of decomposing both data and analytic work enables Q:OS to deliver unprecedented visibility and detection capabilities that continually improve as more data is gradually collected and integrated.

This visibility supports understanding and predictive awareness as Q:OS leverages its deep learning and machine learning algorithms, coupled with mathematical models and statistical analytics. Simulation modeling explores "what if" scenarios and the use of advanced model management tools to improve predictive model performance and management. This exploration of hypothetical strategies and outcomes with continuously enriched data allows CISOs to better understand the intricacies of their entire security posture and deliver a more secure enterprise. On fully instrumented networks, ongoing network resilience scoring that can incorporate asset information, vulnerabilities, exploits, and privilege information better understands the business impact from different incident scenarios. Instrumentation and continuous monitoring of exhaustive endpoint, logging, and other network telemetry data reveal hidden relationships and complex interdependencies that would otherwise be very difficult to recognize or comprehend.

By organizing data into a unified data model upon ingestion data, Q:OS enables databases to evolve into knowledge bases.

Extending these capabilities to other areas of the business, Q:OS can be used to identify overlapping or extraneous workflows and other operational inefficiencies that directly affect the bottom line. By reusing the vast majority of the same infrastructure and technologies to ingest, analyze, enrich, and report on data, the foundational capabilities of Q:OS are extensible to drive visibility and eventually optimization across security, operational, and financial data sets.

QOMPLX:OS

True Competitive Advantage in a Digital World

Q:OS is a hosted and turn-key distributed analytics technology stack that is purpose-built to derive actionable insights from data that can be measured in petabytes. QOMPLX offers a decision platform that is truly an end-to-end solution—as data is collected and aggregated, insights gleaned, and contextualized decisions to act are made, Q:OS monitors these decisions and makes adjustments based on additional data to optimize business growth and human productivity. Leveraging QOMPLX's data management, simulation modeling, and continuous machine learning technologies is a game-changing approach to driving sales and optimizing profitability. Entirely new pillars of business will emerge for companies innovative and insightful enough to embrace the full spectrum of capabilities available in Q:OS.

Key Available Use Cases

- Rapid Data Ingest
- Seamless Data Integration
- Event-Condition-Action (ECA) Automation
- BPM/KPO Optimization
- **Turn-Key Integration with Q:OS Components**
- Ad Hoc Queries
- Advanced Fuzzy Search
- Geospatial Queries
- Distributed Queries
- Advanced Statistical Analytics
- Behavioral Analytics
- Behavior Extraction
- Funnel Analysis
- Lifetime Value of Customers
- Data Source Health Monitoring