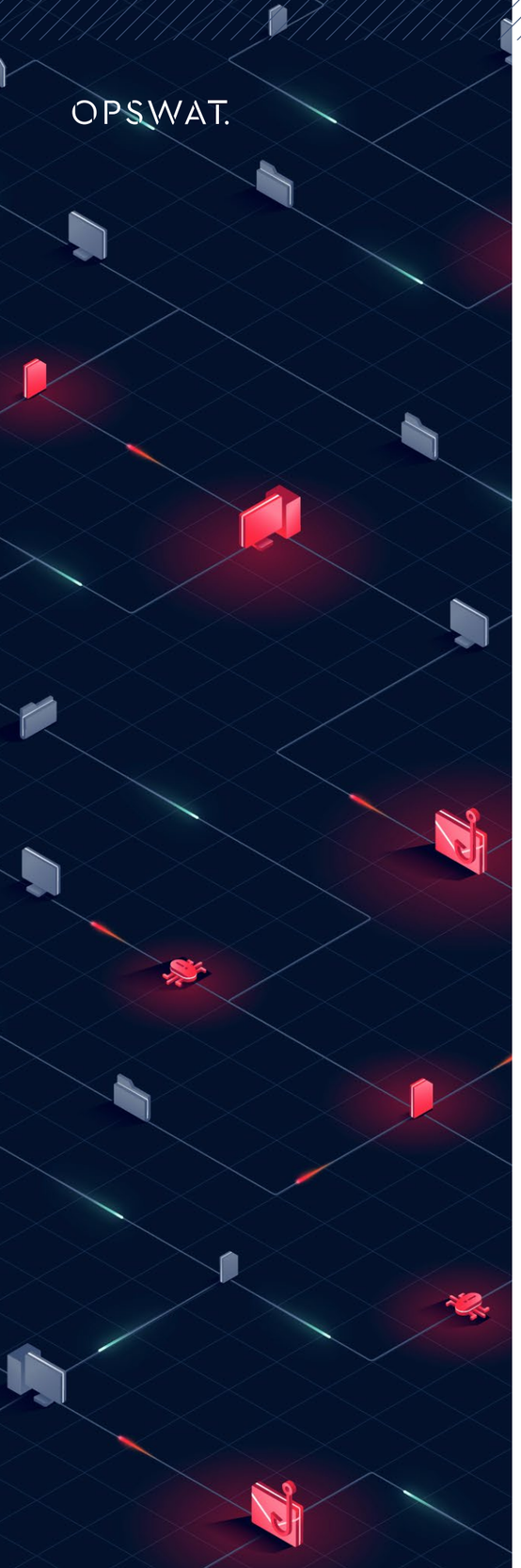
DATA LAKE SECURITY

Protect the Data That Powers AI and Analytics

Ensure only trusted, malware-free, policy-compliant data enters AI models, analytics platforms, and business systems.



Overview

Data lakes have become essential to modern business, powering advanced analytics, machine learning, and data-driven decisions. But their scale and the sensitivity of the data they hold also make them a high-value target. A single breach can expose sensitive information, stall business initiatives, and trigger financial, reputational, and legal fallout.

OPSWAT secures the data lake without slowing the innovation that makes it valuable. Taking a prevention-first approach, the MetaDefender™ Platform inspects and sanitizes data before ingestion, protects it at rest, and secures its movement across trust boundaries, so every analytics and AI decision rests on data you can trust.

Challenges in Securing Data Lakes

Every file, dataset, and third-party feed entering the data lake is a potential attack path. Documents, code, sensor data, and external content flow in continuously, often without inspection, and once a malicious or compromised file is stored, it inherits the trust of everything around it. From there, hidden threats and poisoned data can move straight into the AI models, analytics, and business systems that depend on them.

Untrusted Data	Files, datasets, and third-party content are ingested continuously, often stored and trusted without inspection.
Hidden Malware	Security controls built for network access and permissions leave stored files and datasets unchecked for embedded threats.
AI/Analytics Risk	Poisoned or manipulated data can skew models and dashboards long before anyone traces the problem to its source.

Comprehensive & Consolidated Solution

The OPSWAT MetaDefender Platform offers a prevention-first approach to securing the data that powers AI and analytics. By providing comprehensive protection across the entire data lifecycle, ingestion, storage, and movement,

OPSWAT enables organizations to leverage their data assets with confidence, ensuring that the insights and decisions driving their business are based on secure, validated, and trusted information.

How it Works

01

Inspect and sanitize at ingestion

- Detect known and unknown malware in files and datasets
- Remove hidden and embedded threats before data is trusted
- Verify file authenticity and block disguised or spoofed files
- Flag vulnerabilities, sensitive data, and AI-generated content
- Inspect files inline at the network perimeter before they reach the lake

02

Protect data at rest

- Continuously scan stored data in real time, on demand, or on a schedule
- Detect and remediate threats hidden in data already at rest
- Automatically tag and quarantine risky files
- Keep audit-ready logs of every scan and action

03

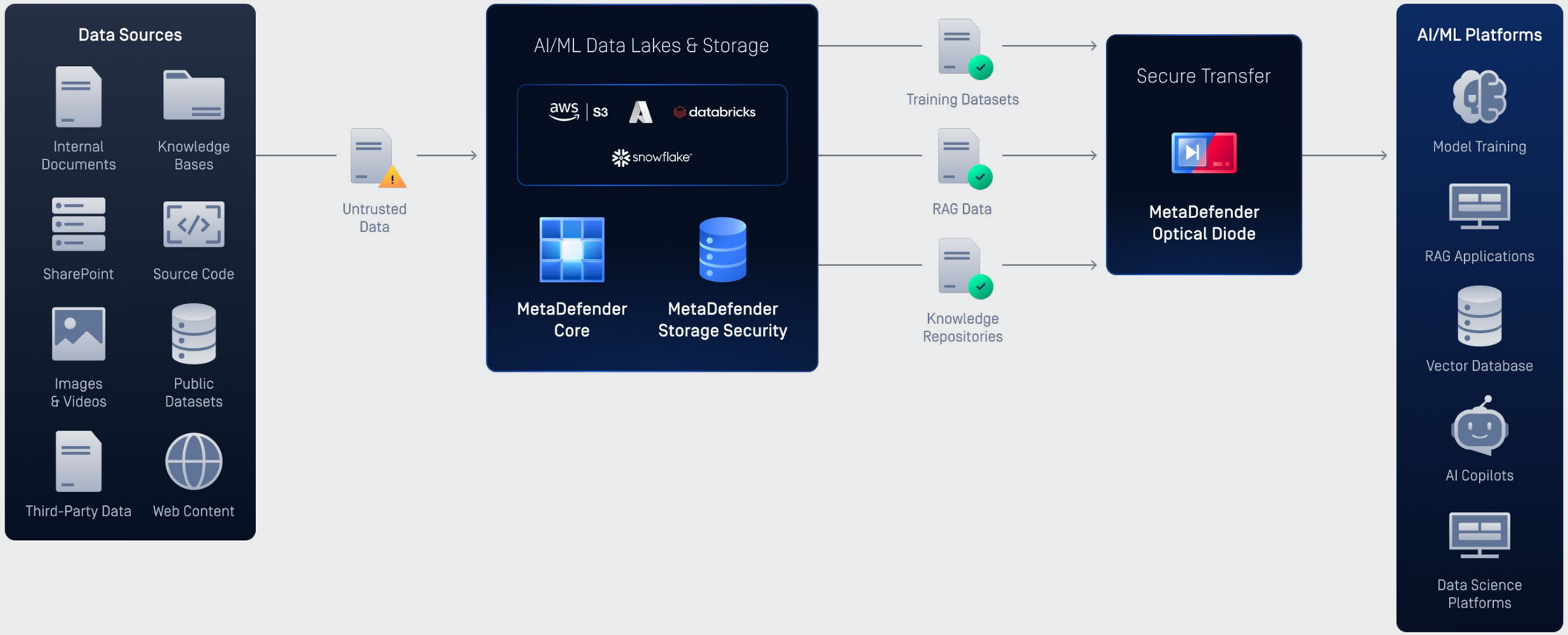
Secure data in motion

- Move approved data securely across environments and trust zones
- Enforce transfer policies and encrypt data in motion
- Automate and route file exchanges without manual handling
- Enforce one-way, air-gapped transfer into isolated or classified environments
- Log every transfer for compliance and audit

One Platform Across Data Lake Lifecycle

AI/ML Training Data

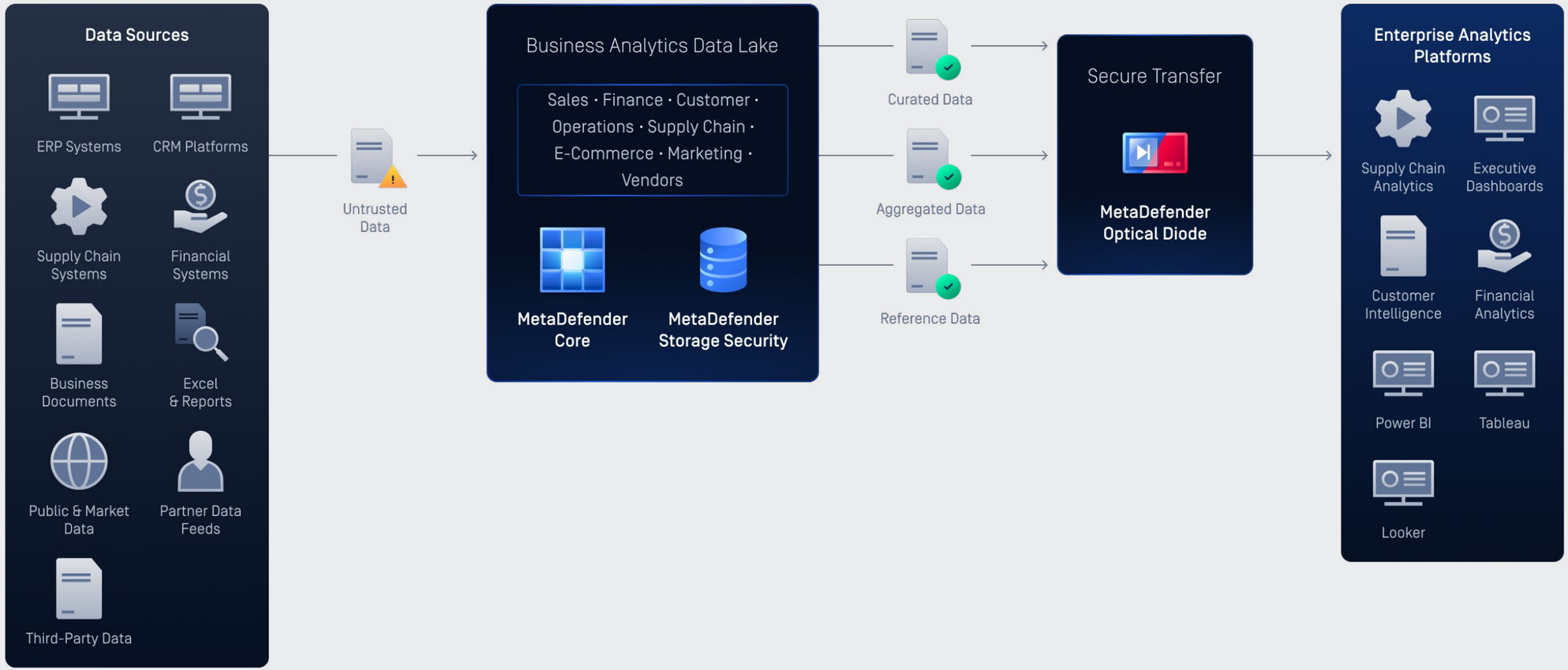
Only trusted documents, code, datasets, and knowledge repositories can enter AI models, LLMs, and RAG applications by keeping poisoned or malicious content out of the data that models learn from.



OPSWAT.

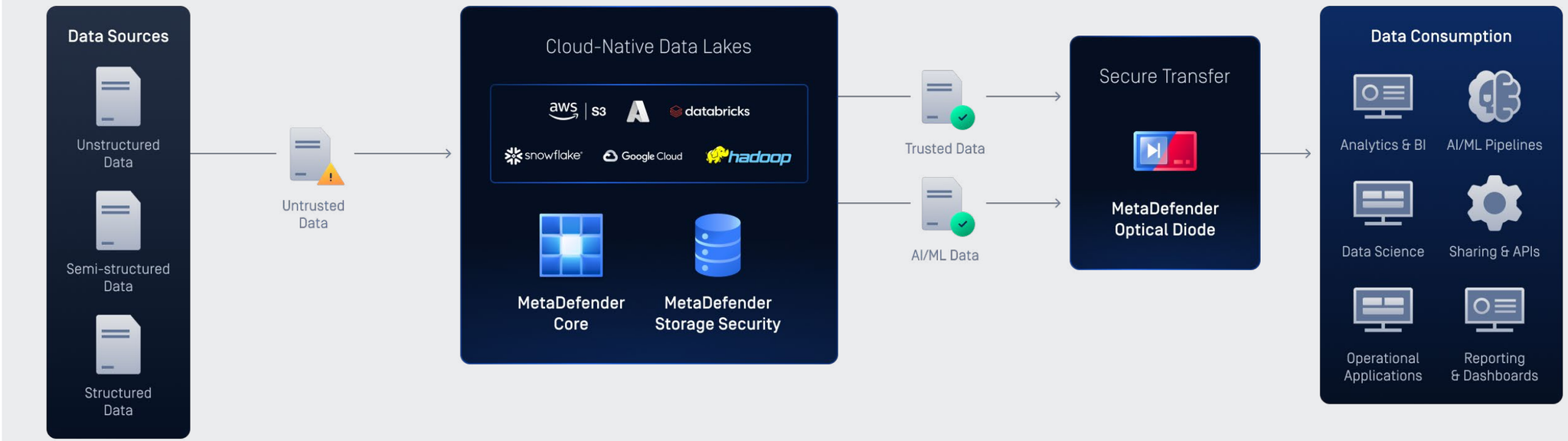
Enterprise Analytics Platforms

Financial, operational, and third-party data feeding dashboards and reports is inspected throughout the analytics lifecycle, so decisions rest on trusted data.



Cloud-Native Data Lakes

Protect cloud-native data lakes across AWS, Azure, and Google Cloud by securing data ingestion, continuously monitoring stored content, and enforcing trust across distributed environments.

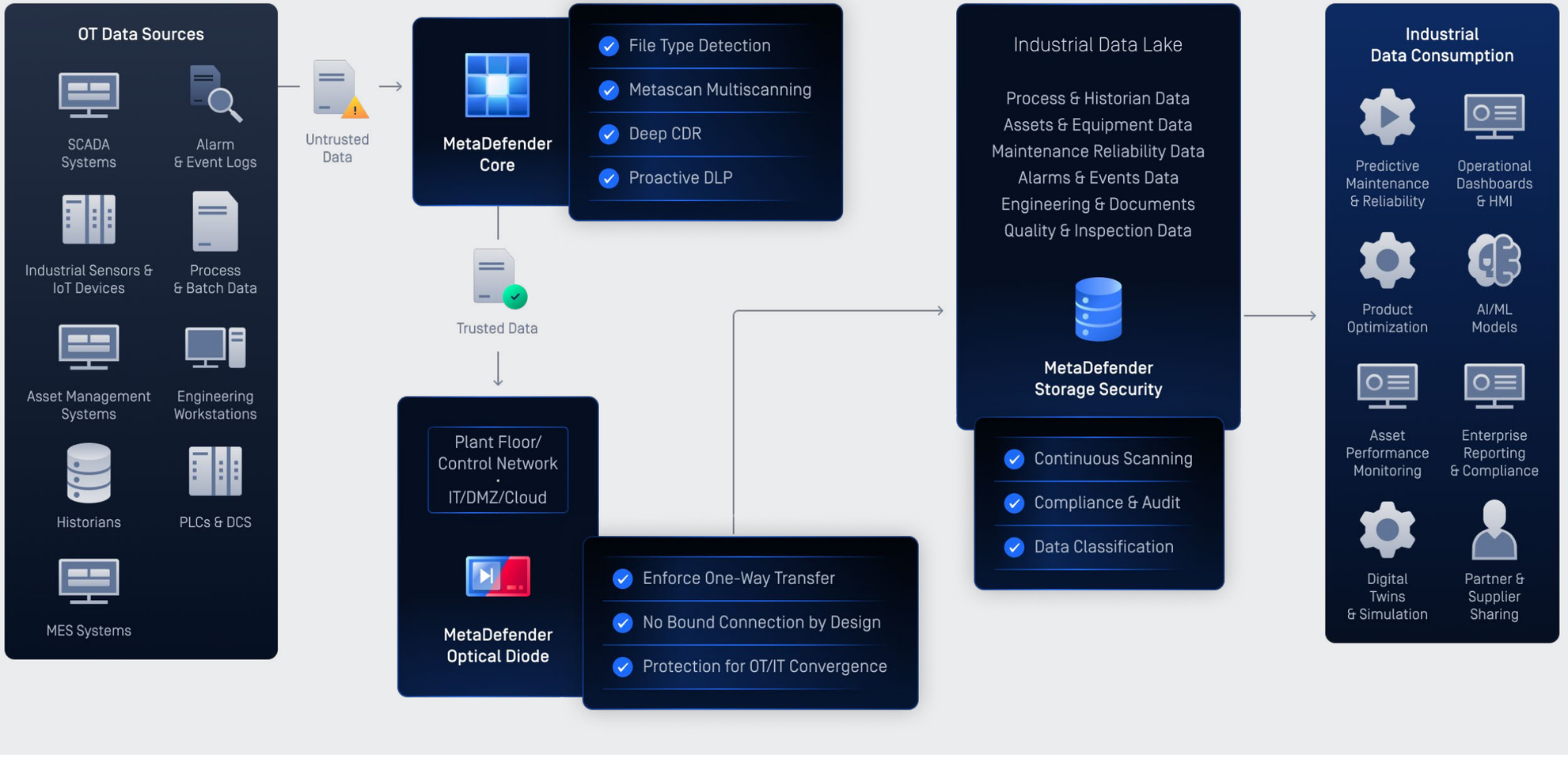


Cross-Domain Data Sharing

Enable secure movement of data across networks, security domains, and trust boundaries, with content validation before transfer and one-way data movement enforced where required.

OT and Industrial Data Lakes

Secure OT data on its journey from industrial environments into data lakes, analytics platforms, and AI systems by inspecting content, enforcing secure one-way transfer, and continuously monitoring stored data.



OPSWAT.

Why OPSWAT

From AI and analytics to cloud, operational technology (OT), and cross-domain environments, OPSWAT helps ensure the data behind critical decisions is secure, validated, and continuously trusted.

Stop Threats at the Door

Inspect and sanitize files, documents, and datasets before they are ingested into the data lake. This prevents malware and weaponized content from becoming part of trusted enterprise data.

Protect Data Continuously

Continuously scan and monitor stored data to detect threats that may have bypassed earlier controls or may have been introduced through internal workflows, ensuring long-term integrity.

Secure Every Handoff

Move approved data into protected AI and analytics environments without exposing critical systems to network-based attacks.

GET STARTED

Trusted data is the foundation every AI and analytics initiative depends on. OPSWAT is here to help you build on solid ground.

Talk to one of our experts today.

Scan the QR code or visit us at:
opswat.com/get-started
sales@opswat.com

OPSWAT.
Protecting the World's Critical Infrastructure

For the last 20 years OPSWAT, a global leader in IT, OT, and ICS critical infrastructure cybersecurity, has continuously evolved an end-to-end solutions platform that gives public and private sector organizations and enterprises spanning Financial Services, Defense, Manufacturing, Energy, Aerospace, and Transportation Systems the critical advantage needed to protect their complex networks from cyberthreats.

Built on a "Trust no file. Trust no device.™" philosophy, OPSWAT solves customers' challenges like hardware scanning to secure the transfer of data, files, and device access with zero-trust solutions and patented technologies across every level of their infrastructure. OPSWAT is trusted globally by more than 2,000 organizations, governments, and institutions across critical infrastructure to help secure their devices, files, and networks from known and unknown threats, zero-day attacks, and malware, while ensuring compliance with industry and government-driven policies and regulations.

Discover how OPSWAT is protecting the world's critical infrastructure and securing our way of life; visit www.opswat.com.