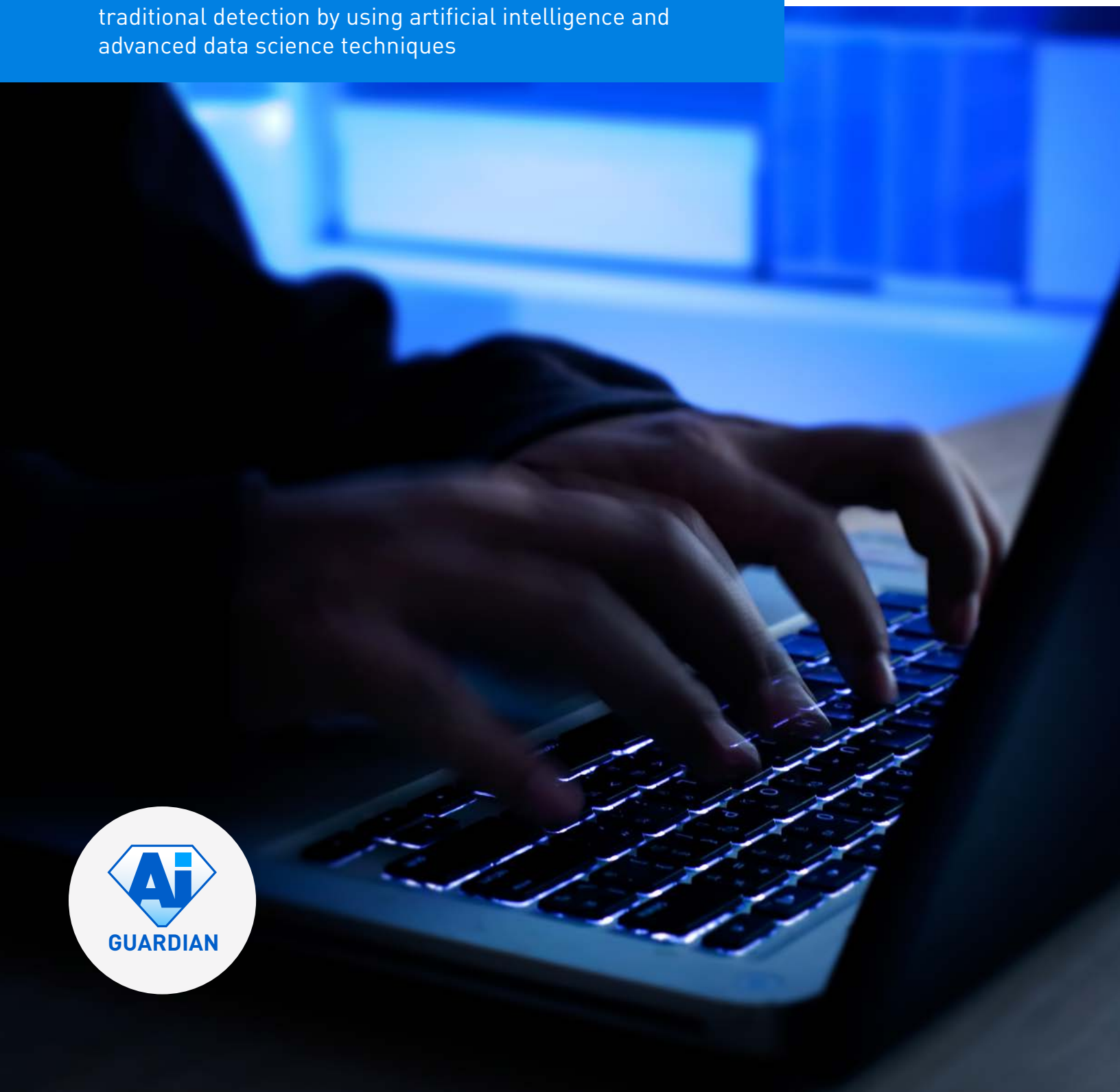
INTERMEDIA®
CLOUD COMMUNICATIONS

WHITEPAPER

AI Guardian

Protecting users from targeted email attacks that evade traditional detection by using artificial intelligence and advanced data science techniques



CONTENTS

EXECUTIVE SUMMARY

Email continues to be the lifeblood of business communications, but it's also become a primary threat vector for cyberattacks that threaten business. As protection against spam, viruses, and known malware become commoditized, cybercriminals have shifted their tactics to craft targeted attacks that exploit human trust and are less obvious threats. Email attacks today are laser focused on specific businesses and users and evade traditional detection by adopting advanced techniques and targeting human nature.

Adversaries research their targets, mask payloads by standing up zero-day domains with redirections, and often impersonate trusted parties to steal money and data. Attackers are also foregoing payloads altogether, focusing instead on socially engineered messages that are expressly crafted to induce certain actions from victims e.g. fraudulently offering iTunes gift cards. Malicious emails are now being delivered from reliable domains such as Gmail and Yahoo to pass authentication checks. Email protection engines based solely on signatures or metadata are no longer enough to protect businesses from these advanced attacks.

AI Guardian from Intermedia takes a data science approach to email protection to stop targeted attacks such as impersonation, payroll fraud, invoice fraud, and zero-day credential phishing. Powered by natural language understanding (NLU), the AI Guardian detection engine analyzes thousands of signals across identity, behavior, language, and global threat data to catch attacks missed by traditional solutions.

This breadth and depth of detection enables AI Guardian to classify threats into predefined categories with high accuracy. These threats are automatically remediated using configurable actions (delete, quarantine) for every threat category. AI Guardian leverages custom machine learning (ML) models for an organization that continually refresh historical baselines and reduce false positives with time.

This paper will summarize current email security challenges and outline how AI Guardian protects against targeted attacks.

THE NEW FACE OF EMAIL ATTACKS

Email attacks are as old as email itself, and attackers have always developed new tactics as security capabilities have matured over the years. While 'click-and-run' attacks like spam and mass phishing campaigns still exist, attackers don't spend too much time crafting them and they can be effectively blocked by traditional security controls.

But cybercriminals have now moved towards email attacks that evade metadata-based detection, don't have binary 'good or bad' payloads, and are finely crafted to push all the right psychological buttons of their intended victims. These attacks - broadly classified under the Business Email Compromise (BEC) umbrella - have dripped and dripped over the years to create a billion dollar ocean. The 2022 IC3 Report from the Federal Bureau of Investigation found that over $43 billion has been lost in BEC attacks over the past five years.

Targeted attacks - both under the BEC category and beyond - usually share these characteristics:

- **Driven by target research:** Targeted attacks avoid the scattergun approach of mass phishing attempts and are the result of extensive groundwork and research conducted by the attacker. The perpetrator is aware of the victim's name, job title, reporting manager, and sometimes even what days they'll be out of office.

- **No malicious payloads:** Targeted attacks rarely include URLs or attachments that contain known malicious payloads, especially in the first email. Payloads may sometimes be introduced at the end of email chains, after the attacker has gained the victim's trust. It's more likely for the 'payload' to be within the email content itself i.e. requests that are framed like they're coming from a legitimate person known to the victim.
- **Rules and metadata are not enough:** Since targeted attacks are more sniper than sledgehammer in their technique, metadata and binary rules are not enough to flag these emails. These protection techniques either lead to a flood of false positives (if they're too strict) or let finely crafted attacks escape their grasp (if they're not strict enough).
- **Socially engineered:** Targeted attacks prey on human nature as much as - if not more than - security controls. Leaning on age-old psychological tricks like urgency, authority, persuasion, and fear, the language in these emails make the victims 'want' to take action without thinking too much about it.

The New Face of Email Attacks

	VECTOR & DELIVERY	TECHNIQUES	PAYLOAD	LEGACY EMAIL CONTROLS
 Spam	Mass email	N/A	Known malicious link or executable	✓
 Mass Phishing	Mass email	Mass-produced phishing kits	Known malicious link or executable	✓
 VIP Impersonation	Gmail/Yahoo, lookalike domains	Social engineering	"Soft" payload as an ask/ request, fake attachments	X
 Payroll Fraud	Gmail/Yahoo, lookalike domains	Impersonation, social engineering	"Soft" payload as an ask/request	X
 Vendor Fraud	Email from compromised account	Impersonation, social engineering	"Soft" payload as an ask/ request, fake attachments	X
 Credential Phishing	Email from compromised account, Gmail/Yahoo	Redirects, brand imperson-ation for login pages, 0-day domains	0-day links, fake attachments	X
 Account Takeover	Credential phishing attack	Auto-forwarding rules, lateral movement	0-day links, fake attachments	X

AI GUARDIAN CAPABILITIES

AI Guardian is a new, additional email security layer for Intermedia Email Protection that protects email communications for Intermedia customers using natural language understanding, artificial intelligence and advanced data science techniques. AI Guardian analyzes thousands of signals across identity, behavior, language, and global threat data. Email threats are classified under predefined detection categories and then automatically remediated (delete, quarantine) based on configurable actions. AI Guardian also detects emails with sensitive PII/PCI information to help organizations stay compliant with industry privacy regulations.*

The tables below outline AI Guardian capabilities.

Inbound Email Protection

CAPABILITIES	DESCRIPTION
Payment Fraud Protection	Detects emails impersonating an external entity to defraud the organization e.g. with fraudulent invoices
Payroll Fraud Protection	Detects emails impersonating an employee to steal money or payroll-related information e.g. with fraudulent W-2 or direct deposit requests
Impersonation Protection	Specific protection against impersonation attacks on VIPs and other key internal staff
Credential Phishing Protection	Detects emails containing links or redirects to fake login pages attempting to steal account credentials
Extortion Protection	Detects emails that threaten users with bad consequences unless they take a specific action
Ransomware Protection	Detects emails that are trying to get users to download and install Ransomware by directing them to access links or to open attachments
Protection Against Socially Engineered Attacks	Detects emails that try to compromise a user's trust to extract information or money from them – e.g. attacks such as benefactor fraud
Protection Against Malicious URLs in Attachments	Scan emails for attachments with URLs that lead to malicious websites

Outbound Email Protection*

Email DLP - PII	Detects the presence of sensitive personally identifiable information (PII) in emails e.g. SSN, passport number etc
Email DLP - PCI	Detects the presence of bank account details, credit card numbers, and other sensitive financial information in emails

*Data Loss Protection (DLP) for PII and PCI data is not available for Microsoft 365 email.

END USER EXPERIENCE

AI Guardian informs and educates users in real time as they encounter suspicious emails.

Labels on suspicious emails
Suspicious emails can be labeled so that users can clearly identify which emails in their inbox are dangerous.

Quarantine folder for suspicious emails
Depending on perceived risk, emails can be quarantined in a separate folder as well, so that end users do not accidentally respond to those emails. Just as in the case of spam folders, users can go into the quarantine folder if they are looking for something specific, but otherwise these folders provide a layer of protection that prevents users from responding to suspicious emails.

Tags on suspicious emails*
Security awareness training has limited value. Users get desensitized to simulated phishing awareness campaigns, or even worse, start alerting each other to simulated attacks, defeating the purpose of the awareness campaign. Tags that teach the user why a suspicious email they legitimately got in their inbox is bad is a much better approach. Body tags help users understand the signs they need to be aware of when they look at emails to inspect their veracity.

LANGUAGE-POWERED DETECTION ENGINE

Targeted email attacks usually don't have one clear red flag - making a determination on these attacks needs a confluence of signals, detection techniques, and data sources.

Detection Signals

The AI Guardian detection engine uses a broad spectrum of signals that span:

- **Identity:** Email security controls need to exhaustively analyze who users are in order to prevent impersonation and spoofing attempts. AI Guardian analyzes users' names, roles, commonly used browsers and clients, email aliases, and so on.
- **Behavior:** Identity is a critical part of email analysis, but these signals can turn noisy if used in isolation. It's also important to analyze what users do, create a behavior baseline, and study any anomalies from this baseline to accurately detect targeted attacks and lateral movement. AI Guardian analyzes communication patterns a user has with internal and external recipients, common domains a company interacts with, common locations and IP addresses that users log in from, and so on.

*Tagging is not available for Microsoft 365 email.

- **Language:** If cybercriminals are able to mask their identity and/or behavior, understanding the language in the email and the intent behind the email can identify signals that stop a pernicious attack. AI Guardian analyzes the language of emails - including attachments - to identify unusual requests made in emails, tones of urgency or fear, whether the email deals with financial topics, and so on.

- **Global threat data:** AI Guardian natively integrates with several threat feeds for real-time threat information. AI Guardian also leverages anonymized data from a global fraud model, enabling it to capture attack learnings from one customer and apply them to every customer.

Detection Techniques

AI Guardian utilizes a broad spectrum of both classical and cutting-edge detection techniques to stop targeted email attacks. As an additional, sophisticated defensive layer along with Intermedia Email Protection, AI Guardian detection techniques enable a more holistic umbrella of protection against varied email attacks.

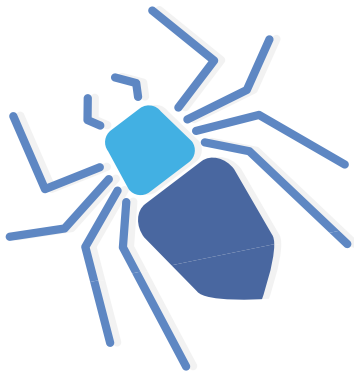
- **Statistical techniques:** Clustering and anomaly detection help identify behaviors outside established norms (e.g. unusual or anonymous IP address logins outside of home and work locations).

- **Traditional machine learning:** Multi-modal classifiers categorize emails into specific buckets (e.g. differentiating between a marketing email, a payroll-related email, and an email with an invoice).

- **Deep learning:** Recurrent neural networks and Long Short Term Memory (LSTM) learn from training data to predict fraudulent and suspicious emails.

- **Natural language understanding:** Entity recognition, sentiment/tone analysis, coreference resolution, and semantic role labeling enable context-aware attack detection.

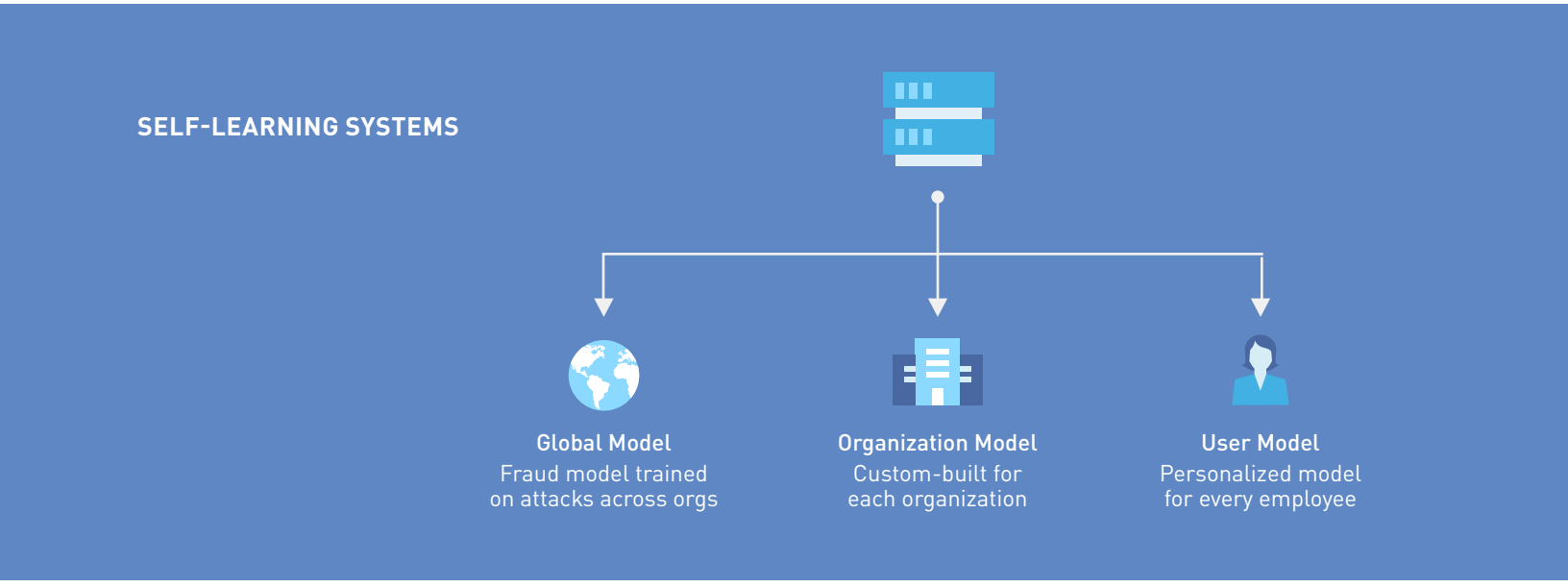
- **Image analysis:** Scans for suspicious invoices and fake login screens using image analysis and computer vision techniques.



SELF-LEARNING SYSTEMS

As long as email remains a critical vehicle for communication, attackers will try and evolve their techniques to bypass known security measures. Ripping and replacing the entire email security stack every few years is not an efficient approach to security. AI Guardian leverages custom machine learning models for every organization for an enterprise-relevant threat detection approach that continuously improves.

- **Learning across organizations:** AI Guardian utilizes anonymized signals across organizations as training data for its global machine learning and fraud prediction models to offer broad and forward-looking email threat protection. Some BEC attacks start with one industry (for example, financial services or local government) and are than replicated across other industries. A model that learns across organizations and industries minimizes this attacker advantage.
- **Learning within organizations:** If learning across organizations offers breadth, building custom self-learning models for each organization offers depth. AI Guardian utilizes custom models that account for the volume and nature of email interactions, frequency of communication across departments, legitimate third-party vendor context, and other data points to provide high-fidelity email threat detection.
- **User-focused learning:** The most focused and possibly deepest level of learning comes from studying individual user identity, behavior, and language signals. AI Guardian builds models for every user. These models take into account topics users discuss, their common login locations, the people they frequently communicate with, their writing styles, and other user-specific signals.



SECURITY AND PRIVACY

The AI Guardian platform processes and remediates suspicious emails in real time while safeguarding your organization's privacy.

Predefined Detection Categories
AI Guardian provides several out of the box detection categories and automatically classifies threats under these categories (e.g. payroll fraud, payment fraud, credential phishing, social engineering etc.). For customers with access to the AI Guardian dashboard, predefined detection categories eliminate the need for custom policy setup and ongoing maintenance.

Automated Remediation
Detected threats can be automatically remediated (e.g. delete, quarantine) by AI Guardian based on configurable remediation actions. Threat remediation actions can be stacked on top of each other and allow for customization according to threat category, user roles (AD), and exceptions.

Data Controls
AI Guardian does not store any customer emails in its cloud, nor does any human look at the email content. Several measures have been put in place to protect the integrity of customer data and ensure that only legitimate services can access the data. AI Guardian models do not contain any information to successfully reverse engineer emails. For more details on AI Guardian data and privacy controls, please get in touch with your Intermedia representative.

SOC2 Type 2 Certified
The AI Guardian platform has been certified under SOC2 Type 2 guidelines, highlighting a commitment to upholding the integrity, confidentiality, and privacy of customer data.

RBAC and Audit Logging
Full role-based access control (RBAC) and detailed audit logs provide visibility into activity.

Access Control
Built-in two-factor authentication (2FA) and single sign-on (SSO) add extra layers of security and protect against compromise.

Readiness for GDPR compliance

In-Memory Pipeline
In-memory processing of the data pipeline reduces copies of data to disks even when data is in transit. This helps maintain integrity while reducing attack surfaces.

Support for Art. 17
Easy APIs to remove identifiable information from datastore within a reasonable timeframe.

Data Monitoring
Proactive monitoring of data access, movement and deletion to ensure that all data within the system is monitored using Prometheus infrastructure.

Security Controls
Enterprise-grade features such as RBAC, AES encryption and MFA help ensure appropriate security controls.

Breach Reporting
Real time alerts to spot unusual activities across the AI Guardian platform help identify, assist and report the breadth of a breach within a short time.

Data Retention
AI Guardian does not persist any non-malicious email content. Data retention within databases can be managed using admin APIs.

BENEFITS

The main benefits of AI Guardian are listed below.

Stop Targeted Email Attacks: Safeguard your business against targeted attacks such as payroll fraud, payment fraud, impersonation, and zero-day credential phishing.

Ensure Regulatory Compliance: Uphold and enforce mandates to protect your organization's sensitive data (SSNs, bank account details, and other PII/PCI data).*

Accelerate Detection & Response: Reduce burden on your security team with predefined detection categories and automated remediation of email threats.

Increase Team Productivity & Happiness: Avoid resource strain with a solution that's easy to deploy, manage, and use. Enable your security team to focus on threats that matter and do more stimulating work by leveraging automation for repeatable tasks.

Protect Brand Reputation: Avoid reputational damage stemming from the loss of sensitive data, paid ransoms, and other negative publicity stemming from security breaches.

Compounding ROI: Leverage email protection that gets smarter every second by learning from attack trends across other organizations as well as from contextual information within your organization.

*Data Loss Protection (DLP) for PII and PCI data is not available for Microsoft 365 email.

Questions? Contact Us Today.
800.379.7729 sales@intermedia.com
intermedia.com