Next-Generation Unified Data Protection Platform

Product Vision: To create an intelligent, adaptive, and seamlessly integrated Data Loss Prevention (DLP) and Rights Management System (RMS) that provides unparalleled data protection across its entire lifecycle, from creation to deletion, regardless of its location or transmission method. This platform will leverage advanced AI/ML capabilities for automated classification and policy enforcement, ensuring sensitive data remains secure and compliant while empowering collaborative work.

Product Backlog: Unified DLP & RMS Solution

Product Goal: Achieve Comprehensive Data Protection & Rights Management

Epics (Major Feature Areas):

- **EPIC 1: Intelligent Data Discovery & Classification**
- **EPIC 2: Advanced Data Loss Prevention (DLP)**
- **EPIC 3: Robust Rights Management System (RMS)**
- **EPIC 4: Seamless Cloud Integration & Hybrid Environment Support**
- **EPIC 5: Auditing, Reporting & Compliance**
- **EPIC 6: User Experience & Administration**
- **EPIC 7: Advanced Security & Threat Intelligence**

Detailed Product Backlog Items (User Stories/Features):

EPIC 1: Intelligent Data Discovery & Classification

- **Story 1.1: Automated Data Discovery (At Rest & In Transit)**
    - As a security administrator, I want the system to automatically discover sensitive data across endpoints, network shares, databases, and cloud storage, so that I have a complete inventory of my organization's data.
    - **Sub-features:**
        - Scheduled and real-time scanning capabilities.
        - Support for various file types (documents, spreadsheets, images, code, etc.).
        - Deep content inspection (OCR for images, analysis of structured/unstructured data).
- **Story 1.2: AI/ML-Infused Automatic Data Classification**

    - As a security administrator, I want the system to automatically classify data based on content, context, and user behavior using AI/ML, so that data protection policies can be applied intelligently without manual intervention.
    - **Sub-features:**
        - Pre-defined classification categories (e.g., PII, PCI, HIPAA, Confidential, Internal).
        - Customizable classification categories.
        - Machine learning models for pattern recognition, natural language processing (NLP), and entity recognition.
        - Confidence scoring for classifications.
        - Ability to learn from manual classification corrections.
- **Story 1.3: Contextual Data Classification**
    - As a security administrator, I want the system to classify data not just by content, but also by its source, creator, last modifier, and location, so that policies are more nuanced and effective.
- **Story 1.4: Auto-Assignment of Classification Categories**
    - As a security administrator, I want the system to automatically assign appropriate data classification categories upon creation or modification, so that data is protected from its inception.
- **Story 1.5: Version Control for Classified Data**
    - As a user, I want the system to maintain version control for classified data, so that I can revert to previous versions and track changes to sensitive information.
    - As a security administrator, I want to track changes and access history of all versions of classified data.
- **Story 1.6: Manual Override & Review for Classification**
    - As a security administrator, I want to be able to manually review and override AI/ML-assigned classifications, so that I can ensure accuracy and compliance.

EPIC 2: Advanced Data Loss Prevention (DLP)

- **Story 2.1: Data at Rest DLP**
    - As a security administrator, I want to define and enforce policies for sensitive data stored on endpoints, network shares, and databases, so that unauthorized access or exfiltration is prevented.
    - **Sub-features:**
        - Quarantine, encryption, or deletion of non-compliant data.
        - Alerting on policy violations.
        - Remediation actions (e.g., block, encrypt, notify).

- **Story 2.2: Data in Transit DLP**
    - As a security administrator, I want to monitor and control the transmission of sensitive data across various communication channels, so that data leakage is prevented during sharing.
    - **Sub-features:**
        - **Email Monitoring & Blocking:** Inspect outgoing emails for sensitive content and block or encrypt as per policy.
        - **Web Upload Monitoring & Blocking:** Prevent sensitive data from being uploaded to unauthorized websites.
        - **Instant Messaging & Collaboration Tools:** Monitor and control data sharing on platforms like Slack, Microsoft Teams, etc.
        - **SFTP/FTP Monitoring & Control:** Enforce policies for data transfers via secure and insecure file transfer protocols.
        - **USB/Removable Media Control:** Prevent unauthorized copying of sensitive data to external devices.
        - **Network Protocol Inspection:** Analyze data traversing common network protocols (HTTP, HTTPS, FTP, SMB, etc.).
- **Story 2.3: Policy Enforcement based on Classification**
    - As a security administrator, I want to create DLP policies that automatically apply based on the assigned data classification, so that protection is consistent and scalable.
- **Story 2.4: Dynamic Policy Adjustment (Contextual DLP)**
    - As a security administrator, I want the system to dynamically adjust DLP policies based on user role, location, device, and network, so that protection is adaptive to real-world scenarios.
- **Story 2.5: User Education & Justification Prompts**
    - As a user, when I am about to perform an action that violates a DLP policy, I want to receive a prompt explaining the violation and requesting justification, so that I understand the policy and can proceed responsibly (or be blocked).
    - As a security administrator, I want to review user justifications for policy overrides.

EPIC 3: Robust Rights Management System (RMS)

- **Story 3.1: Persistent Data Encryption**
    - As a security administrator, I want the system to encrypt sensitive data persistently, regardless of where it is stored or who accesses it, so that data remains protected even if exfiltrated.
- **Story 3.2: Granular Access Control**

    - As a security administrator, I want to define granular access permissions (e.g., view, edit, print, copy, forward, screenshot) for classified data, so that only authorized users can perform specific actions.

- **Story 3.3: Dynamic Rights Revocation**
    - As a security administrator, I want to be able to revoke access rights to data dynamically, even after it has been shared, so that control over sensitive information is maintained at all times.
- **Story 3.4: Time-Based Access Control**
    - As a security administrator, I want to set time-based access limitations for shared data, so that access automatically expires after a defined period.
- **Story 3.5: Watermarking & Visual Deterrents**
    - As a security administrator, I want to apply dynamic watermarks to sensitive documents (e.g., user name, IP address, timestamp) when viewed or printed, so that unauthorized sharing is deterred and traceable.
- **Story 3.6: Offline Access with Policy Enforcement**
    - As a user, I want to be able to access RMS-protected data offline, with all defined rights and policies still enforced, so that I can work productively without constant internet connectivity while maintaining security.
- **Story 3.7: Secure Sharing with External Parties**
    - As a user, I want to securely share classified data with external partners, with all RMS policies enforced for them, so that collaboration is secure.
    - As a security administrator, I want to manage and monitor external party access to classified data.

EPIC 4: Seamless Cloud Integration & Hybrid Environment Support

- **Story 4.1: Integration with OneDrive**
    - As a security administrator, I want the system to seamlessly integrate with OneDrive for data discovery, classification, DLP, and RMS policy enforcement, so that data in OneDrive is protected.
- **Story 4.2: Integration with Google Drive**
    - As a security administrator, I want the system to seamlessly integrate with Google Drive for data discovery, classification, DLP, and RMS policy enforcement, so that data in Google Drive is protected.
- **Story 4.3: Integration with Dropbox**
    - As a security administrator, I want the system to seamlessly integrate with Dropbox for data discovery, classification, DLP, and RMS policy enforcement, so that data in Dropbox is protected.

- **Story 4.4: Integration with Other Cloud Storage Providers (e.g., Box, SharePoint Online, AWS S3, Azure Blob Storage)**
    - As a security administrator, I want the system to offer easy integration with other popular cloud storage services, so that our entire cloud footprint can be secured.
    - **Considerations:** API-based integration, proxy-based integration.
- **Story 4.5: On-Premise & Hybrid Environment Support**
    - As a security administrator, I want the solution to effectively protect data in both on-premise and hybrid cloud environments, so that consistent policies can be applied across our entire infrastructure.
- **Story 4.6: Cloud-to-Cloud DLP**
    - As a security administrator, I want to prevent sensitive data from being moved between unauthorized cloud storage providers, so that data remains within approved environments.

EPIC 5: Auditing, Reporting & Compliance

- **Story 5.1: Comprehensive Activity Logging**
    - As a security administrator, I want the system to log all data access, modification, sharing, and policy violation attempts, so that I have a complete audit trail.
- **Story 5.2: Customizable Reporting & Dashboards**
    - As a security administrator, I want customizable reports and interactive dashboards that provide insights into data classification, DLP incidents, RMS usage, and compliance status, so that I can quickly identify risks and demonstrate compliance.
    - **Sub-features:**
        - Pre-built compliance reports (GDPR, HIPAA, PCI DSS, etc.).
        - Trend analysis and anomaly detection in data usage.
- **Story 5.3: Real-time Alerts & Notifications**
    - As a security administrator, I want to receive real-time alerts for critical policy violations or suspicious activities, so that I can respond immediately to potential data breaches.
- **Story 5.4: Forensic Analysis Capabilities**
    - As a security administrator, I want the ability to drill down into specific incidents for forensic analysis, including who accessed what data, when, from where, and what actions were taken, so that I can conduct thorough investigations.
- **Story 5.5: Integration with SIEM/SOAR Systems**
    - As a security administrator, I want the system to seamlessly integrate with our existing SIEM (Security Information and Event Management) and SOAR (Security Orchestration, Automation and Response) platforms, so that security events can be centrally managed and automated responses triggered.

EPIC 6: User Experience & Administration

- **Story 6.1: Intuitive Administrative Console**
    - As a security administrator, I want a user-friendly and intuitive web-based console for managing policies, monitoring incidents, and generating reports, so that I can efficiently manage the system.
- **Story 6.2: Role-Based Access Control (RBAC)**
    - As a security administrator, I want to define different administrative roles with specific permissions, so that responsibilities can be delegated securely.
- **Story 6.3: Seamless User Experience (Minimal Disruption)**
    - As an end-user, I want the DLP and RMS features to operate in the background with minimal disruption to my workflow, so that productivity is not hindered.
    - **Consideration:** Agent-based deployment, transparent encryption.
- **Story 6.4: Multi-Tenancy Support**
    - As a service provider, I want the system to support multi-tenancy, so that I can offer DLP/RMS services to multiple organizations with isolated environments. (Optional, if targeting MSPs).
- **Story 6.5: API for Custom Integrations**
    - As a developer, I want a well-documented API to integrate the DLP/RMS solution with other enterprise applications, so that custom workflows and automation can be created.

EPIC 7: Advanced Security & Threat Intelligence

- **Story 7.1: Behavioral Analytics for Insider Threat Detection**
    - As a security administrator, I want the system to analyze user behavior patterns to detect anomalous activities that might indicate insider threats, so that proactive measures can be taken.
- **Story 7.2: Integration with Threat Intelligence Feeds**
    - As a security administrator, I want the system to integrate with external threat intelligence feeds, so that it can identify and block communication with known malicious domains or IPs when sensitive data is involved.
- **Story 7.3: Data Anonymization/Pseudonymization (Optional but highly valuable)**
    - As a data privacy officer, I want the system to be able to anonymize or pseudonymize sensitive data in test or development environments while maintaining data utility, so that compliance with privacy regulations is ensured without hindering development.
- **Story 7.4: Homomorphic Encryption (Advanced - Future Consideration)**
    - As a security researcher, I want the system to explore homomorphic encryption for certain operations, so that data can be processed while remaining encrypted, further enhancing security.

Additional Features to make it one of the best DLP products:

- **Traffic Light Protocol (TLP) or Similar Protocols for Data Sharing:**
    - As a user, when sharing classified data, I want to be able to assign a TLP (e.g., TLP:RED, TLP:AMBER, TLP:GREEN, TLP:WHITE) or a similar internal classification (e.g., "Company Confidential - Do Not Share Externally"), so that the recipient immediately understands the handling restrictions.
    - As a security administrator, I want the system to enforce policies based on the assigned TLP, preventing sharing of "RED" data outside the organization.
    - **Consideration:** This could be integrated as a pre-sharing prompt or a mandatory field for classified data being shared.
- **Blockchain for Data Provenance/Integrity (Future/Advanced):**
    - As a forensic investigator, I want an immutable ledger of all data access and modification events, so that the integrity and provenance of sensitive data can be indisputably verified.
- **Container Security Integration:**
    - As a DevOps engineer, I want the DLP/RMS solution to extend its protection to data within containerized environments (Docker, Kubernetes), so that sensitive data in microservices and cloud-native applications is secured.
- **Secure Print & Scan Control:**
    - As a security administrator, I want to monitor and control what sensitive data can be printed or scanned, and apply watermarks or block actions as per policy.
- **Virtual Desktop Infrastructure (VDI) Support:**
    - As a security administrator, I want the DLP/RMS solution to fully support VDI environments, ensuring data protection for remote and virtualized desktops.
- **Offline Policy Enforcement (Enhanced):**
    - Beyond basic access, ensure comprehensive DLP and RMS policies are enforced even when endpoints are completely disconnected from the network for extended periods.
- **Adaptive Security Policies:**
    - Leverage AI/ML to continuously analyze data, user behavior, and threat landscapes to proactively suggest and adapt security policies to emerging risks.
- **Integration with Enterprise Content Management (ECM) Systems:**
    - Seamlessly integrate with ECM systems (e.g., SharePoint on-prem, OpenText) to apply DLP/RMS policies to documents within these repositories.

This comprehensive backlog provides a strong foundation for your startups to innovate and build a truly cutting-edge DLP and RMS solution. Encourage them to prioritize, iterate, and gather feedback throughout the development process. Good luck to your startups!