PKWARE

ENCRYPTION WITHOUT COMPROMISE

SMARTCRYPT

TECHNICAL BRIEF: DISCOVERY SCANNING

PKWARE's Smartcrypt integrates intelligent data discovery with strong data-level encryption, allowing organizations to identify sensitive information and protect it against loss, theft or misuse.

Companies can use Smartcrypt to find and protect data based on mandates such as PCI-DSS or HIPAA, and can also scan for intellectual property and other forms of sensitive data.

Policy Definition and Configuration

Smartcrypt uses software agents to scan desktops, laptops, and servers for sensitive information. Agents are controlled by the Smartcrypt Enterprise Manager, which applies the organization's discovery and encryption policies.

When configuring Smartcrypt, administrators define the data types that require protection using "Smart Filter Bundles" composed of the following elements:

- **Patterns**: unique data formats (such as credit card numbers or SSNs)
- **Dictionaries**: lists of search terms (such as "patent" or "confidential")
- **Regular expressions**: administrator-defined logical statements
- **Thresholds**: the number of occurrences of a given data type

Smartcrypt includes three pre-defined bundles: HIPAA, PCI-DSS, and Personally Identifiable Information. Administrators can create additional bundles using the elements listed above.

The Discovery Process

Smartcrypt Data Discovery applies a rigorous sequence of rules in order to minimize the possibility of false positives or false negatives.

- Each time a file is added or modified, Smartcrypt receives a file change notification from the operating system.
- If the new or modified file is in a location Smartcrypt has been configured to protect, the agent initiates a discovery scan.
- The file contents are unpacked and extracted as text (extraction is to device memory rather than to disk).
- The Smartcrypt agent inspects the file contents and applies pattern identification logic (including fuzzy logic) to determine whether the file contains sensitive data as defined in the organization's policies.

Smartcrypt can scan any form of unstructured data that can be extracted to text, including ZIP files, files whose extensions have been modified or removed, and file metadata.

SAMPLE FILTER BUNDLES

The following examples are just a few of the filter bundles that an organization could create in the Smartcrypt Enterprise Manager, depending on its business needs and compliance obligations:

» PCI-DSS - Credit Cards

» PCI-DSS - Credit Cards & IBAN

» PCI-DSS - Credit Cards & postal addresses

» PCI-DSS - Credit Cards and e-mail addresses

» PCI-DSS - Credit Cards with CVV

» PCI-DSS - Credit Cards with CVV and e-mail addresses

» HIPAA - Prescription Drugs and Personally Identifiable Information

» HIPAA - ICD-10 & Diagnosis Lexicon

» HIPAA - Personally Identifiable Information

» HIPAA - Pharmaceutical firms, drugs and diagnosis

» HIPAA - Prescription Drugs

» GDPR - Credit Card Numbers and e-mail address

» GDPR - Date of birth and national ID

PKWARE
ENCRYPTION WITHOUT **COMPROMISE**

Discovery Objective	Smartcrypt Data Definition
PCI-DSS Discovery	**Pattern:** Credit Card Numbers: • Visa • MasterCard • American Express • Diners • Discover • JCB **Pattern:** US Social Security Numbers **Pattern:** US Tax IDs **Pattern:** International Bank Account Numbers
HIPAA Discovery	**Dictionary:** FDA Drugs **Dictionary:** FDA Firms **Dictionary:** ICD-9 and ICD-10 Codes **Pattern:** UK National Insurance Numbers **Pattern:** US Social Security Numbers
PII Discovery	**Pattern:** US Addresses (10 or more) **Pattern:** Email Addresses **Pattern:** UK National Insurance Numbers **Pattern:** US Social Security Numbers **Pattern:** Foreign Registration Numbers **Pattern:** US Tax IDs **Pattern:** Government IDs and Passports
Intellectual Property Discovery	**Dictionary:** Intellectual Property Custom dictionary uploads

For more information, including a full list of the terms and phrases included in the dictionaries referenced above, please visit our support site:

https://support.pkware.com/display/SMAR/Distributed+Dictionaries

EXAMPLE: CREDIT CARD NUMBERS

When scanning a file to determine whether it contains credit card numbers, Smartcrypt searches for sequences of numbers that match algorithms and formats published by credit card issuers such as Visa, MasterCard, and American Express.

Smartcrypt can identify valid credit card numbers in any of the following formats:

- 1234123412341234
- 1234-1234-1234-1234
- 1234 1234 1234 1234

The Smartcrypt agent will ignore similar number sequences that are not valid card numbers based on the Luhn test or other relevant algorithms.

EXAMPLE: SOCIAL SECURITY NUMBERS

When configured to scan for Social Security numbers, Smartcrypt uses the guidance published by the Social Security Administration, together with contextual logic and other validation rules, to determine whether a given nine-digit sequence of numbers is a valid SSN.

To avoid false positives, Smartcrypt will only identify a number as an SSN if it is separated by dashes in the correct locations:

111223333: will be ignored
111-223-333: will be ignored
111 22 3333: will be ignored
111-22-3333: will be identified as an SSN

PKWARE®
www.pkware.com

CORPORATE HEADQUARTERS
201 E. Pittsburgh Ave.
Suite 400
Milwaukee, WI 53204

+ 1 866 583 1795

EMEA HEADQUARTERS
79 College Road
Suite 221
Harrow HA1 1BD

+ 44 (0) 203 367 2249