



このコンテンツは公開から3年以上経過しており内容が古い可能性があります  
最新情報については[サービス別資料](#)もしくはサービスのドキュメントをご確認ください

# [AWS Black Belt Online Seminar]

## Amazon Macie



サービスカットシリーズ

Solutions Architect 高橋 悟史  
2020/8/12

AWS 公式 Webinar  
<https://amzn.to/JPWebinar>



過去資料  
<https://amzn.to/JPArchive>



# 自己紹介

名前：高橋 悟史（たかはし さとし）  
CISSP, CISM

職種：セキュリティ ソリューション アーキテクト

職務：AWS セキュリティの啓蒙活動、セキュリティ関連サービスに関するお客様のサポート

経歴：外資ITベンダーにて、メインフレーム、HPC、分散系システムの構築運用に関わったあと、システム管理、セキュリティ、特に認証やID管理、SSOに関する技術支援に従事。セキュリティ企業にてサイバーセキュリティソリューション、SIEMを担当。SaaSベンダーにてクラウド・セキュリティを担当したのちにAWSへ

好きな AWS サービス：AWS Key Management Service



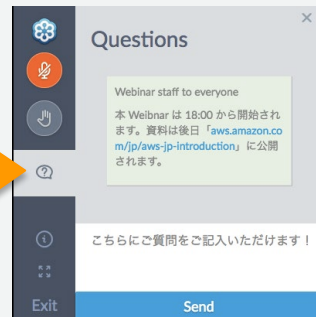
# AWS Black Belt Online Seminar とは

「サービス別」「ソリューション別」「業種別」のそれぞれのテーマに分かれて、アマゾンウェブ サービス ジャパン株式会社が主催するオンラインセミナーシリーズです。

## 質問を投げることができます！

- 書き込んだ質問は、主催者にしか見えません
- 今後のロードマップに関するご質問は  
お答えできませんのでご了承下さい

- ① 吹き出しをクリック
- ② 質問を入力
- ③ Sendをクリック



Twitter ハッシュタグは以下をご利用ください  
#awsblackbelt

# 内容についての注意点

- 本資料では2020年8月12日時点のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(<http://aws.amazon.com>)にてご確認ください。
- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。
- 価格は税抜表記となっております。日本居住者のお客様には別途消費税をご請求させていただきます。
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

# アジェンダ

- 機微情報管理の重要性と課題
- Amazon Macie 概要
- Amazon S3 バケットとオブジェクトの使用状況の可視化
- 機微情報の検出・評価
- 評価・検出結果の参照と、他のサービスとの連携
- 管理・運用・制約

# 機微情報管理の重要性と難しさ

機微情報とは

個人情報、クレジットカード、銀行口座番号などの保護が必要な情報  
特許情報、設計技術など競合他社に知られてはいけない情報

- 機微情報は、既に、認証、アクセス制御、暗号化、ログ管理などで保護している
- 社内の情報規定で対策されているにも関わらず、気づかないところに機微情報が存在したら？

- データレイク上の保管データの中に機微情報が紛れてしまっていたら？
- ログデータの中に意図せず個人情報が入ってしまっていたら？

機微情報が正しく管理され保護されていると自信を持って言えるか？

# Amazon Macie が解決する課題

1

アカウント内にある Amazon S3 バケットの利用状況の可視化、格納されている大量のオブジェクトの可視化

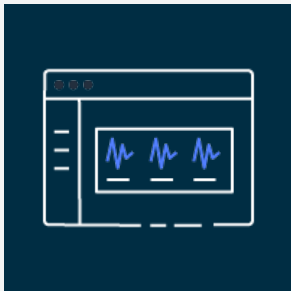
2

設定に基づき、指定されたバケット内の機微情報の評価・検出を効率的に実行

3

評価・検出結果の参照、他のサービスへの連携

# Amazon Macie 概要



Amazon S3 バケットの使用状況の可視化と評価

- バケットの使用状況とバケットポリシーの可視化



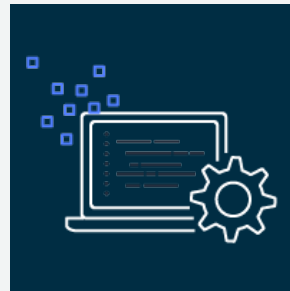
機微情報の検出

- 検出ジョブによる検出
- 柔軟な検出範囲と条件の設定
- AWS マネージド定義とカスタム定義の両方を利用可能



スケーラビリティと集中管理の両立

- AWS Organizations 対応



検出結果に対する対応の実行と自動化

- 詳細な結果 (Findings) の出力とイベントによる AWS サービスとの連携
- 管理 API の提供



# 1. Amazon S3 バケットとオブジェクト の使用状況の可視化

# Amazon Macie の利用ステップ

- Amazon Macie では、アカウントが持っている S3 バケットを全部スキャンするのではなく、管理者が S3 バケットの状態を確認して、スキャンが必要であるバケットを指定してスキャンする仕組みになっている
- AWS マネジメントコンソール上の Macie の管理画面で S3 バケットの状況サマリー、それぞれのバケットの詳細を確認して、スキャン対象のバケットを指定する
- スキャンするためにジョブを作成する。ジョブはスキャン対象バケットなどの対象の設定とスケジュールがセットになっている。スケジュールした場合、指定したスケジュールで自動でジョブが実行される

# Amazon S3 バケットの概要表示 1

AWS マネジメントコンソール→Amazon Macie→概要  
S3 バケットの利用状況を把握可能

バケット数、ストレージ利用量、オブジェクト数

S3 バケットの概要 [詳細はこちら](#)

最終更新日: 05-15-2020 09:00:47

S3 バケットの合計数

12

合計ストレージ

3.92 GB

オブジェクト数

1.15m

アカウント

## パブリック

0%

バケットの 0% はパブリックアクセス可能です

パブリックアクセス可能

0

パブリックアクセスできない

12

世界中パブリックに書き込み可能

0

世界中パブリックに読み取り可能

0

バケット公開状況

## 暗号化されていない

92%

バケットの 92% は暗号化されていません

暗号化されていない

11

暗号化済み

1

SSE-S3 で暗号化済み

1

SSE-KMS で暗号化済み

0

バケット暗号化状況

## 共有済み

0%

バケットの 0% は組織のアカウント外で共有されています

外部で共有されています

0

外部では共有されていません

12

内部で共有されています

0

共有されていません

12

バケット共有状況

# Amazon S3 バケットの概要表示 2

前ページのスクリーンショットの下側の画面  
S3 バケットの設定変更を Findings として検出する  
これは Macie の機微情報のスキャンとは別機能となる

| Top S3 buckets<br>Past 7 days                |         | Top finding types<br>Past 7 days          |         |
|--|---------|---|---------|
| S3 Bucket                                    | 検出結果の合計 | Finding type                              | 検出結果の合計 |
| create-sample-finding-bucket                 | 1       | Policy:IAMUser/S3BucketEncryptionDisabled | 1       |
| <a href="#">View all findings by buckets</a> |         | <a href="#">View all findings by type</a> |         |

| 最新のポリシー検出結果 <a href="#">詳細はこちら</a>                            | 🔄  |
|---|----|
| Most recent policy findings                                   |    |
| <span>Medium</span> Policy:IAMUser/S3BucketEncryptionDisabled | 秒前 |

例 : S3 バケットの暗号化が無効化されたことを検出

# Amazon S3 バケットごとの状況表示

Amazon Macie → S3 バケット でアカウントが持つバケットの状況を個別に参照出来る

一覧表示するバケットをフィルタ可能  
(次ページで説明)

The screenshot shows the Amazon S3 console interface. On the left, a list of buckets is displayed with columns for bucket name, size, and object count. The bucket 'labdatrailbucket' is highlighted with a red box. A red arrow points from this bucket to a detailed view on the right. The detailed view shows various settings for the bucket, including account ID, region, ARN, name, creation date, object count, access type, size, compression size, versioning, last analysis date, encryption type, public access, replication, and tags.

| バケット名            | サイズ    | オブジェクト数 |
|------------------|--------|---------|
| aws-athena-...   | 457 MB | 36      |
| aws-athena-...   | 23 MB  | 36      |
| aws-glue-scr...  | 1 MB   | 6       |
| aws-glue-ter...  | 1 MB   | 1       |
| config-bucke...  | 41 MB  | 2,24k   |
| labdatrailbui... | 1 GB   | 478,68k |
| my-siem-s3-      | 599 MB | 18,24k  |
| satoshi-inter... | 2 MB   | 22      |
| satoshi-tech     | 0      | 0       |
| satstak-flow     | 20 MB  | 30,87k  |
| satstak-inter    | 1 GB   | 621,94k |
| satstak-shar     | 54 KB  | 2       |

| 概要          |                               |
|-------------|-------------------------------|
| アカウント ID    | [Redacted]                    |
| リージョン       | アジアパシフィック (東京) ap-northeast-1 |
| ARN         | arn:aws:s3::labdatrailbucket  |
| 名前          | labdatrailbucket              |
| 作成済み        | 04-10-2019 15:17:00 (1 年前)    |
| オブジェクト      | 478682                        |
| 分類可能なオブジェクト | 478682                        |
| 共有アクセス      | Not shared                    |
| サイズ         | 1 GB                          |
| 圧縮サイズ       | 1 GB                          |
| バージョンング     | 無効                            |
| 最終分析日       | 05-15-2020 09:00:46 (2 時間 前)  |

| Number of objects by encryption type |        |
|--------------------------------------|--------|
| カスタマーマネージド                           | 0      |
| SSE-KMS マネージド                        | 0      |
| SSE-S3 マネージド                         | 478682 |
| 暗号化なし                                | 0      |

| Public access |            |
|---------------|------------|
| 有効なアクセス許可     | Not public |

| Replication   |     |
|---------------|-----|
| レプリケート済み      | いいえ |
| 外部でレプリケート済み   | いいえ |
| レプリケーションアカウント | なし  |

| Tags |    |
|------|----|
| タグ   | なし |

| アカウントレベル 設定         |    |
|---------------------|----|
| ACL によるパブリックアクセスをプロ | オン |
| バケットポリシーによるパブリックア   | オン |

左のペインにバケットの一覧が表示され、バケットをクリックすると右側に詳細が表示される

- オブジェクト数、暗号化の有無と暗号化方式
- 公開（パブリック・アクセス）の有無
- レプリケーション設定の有無

# Amazon S3 バケットのフィルタ表示

## バケットのフィルタで指定出来る条件

- アカウントID
- オブジェクト数
- ストレージサイズ
- タグキー
- バージョニング有無
- パブリック・アクセス許可の有無（書き込み、読み込み）
- レプリケートの有無
- 暗号化方式別のオブジェクト数（SSE-KMS(AWS), SSE-KMS(Customer), SSE-S3, 暗号化無し）
- バケットの暗号化設定

## ユースケース

- オブジェクト数が多いバケットを選択する
- 暗号化されていないオブジェクトがあるバケットを選択する
- リスクの高いバケットを優先的にスキャンすることが可能

# ジョブの作成

対象とするバケットにチェックを入れて右上のジョブを作成ボタンを押すと、ジョブ作成の画面が表示される

The screenshot shows the AWS S3 console interface. At the top, there is a search bar and a button labeled "ジョブを作成" (Create Job) which is highlighted with a red box. Below this is a table of S3 buckets. The bucket "satstak-macie-test" is selected, indicated by a blue checkmark and a red box around the row. To the right of the bucket list is a sidebar for the selected bucket, titled "satstak-macie-test". This sidebar contains various details about the bucket, including its account ID, region, ARN, name, creation date, object count, and public access settings. The "Number of objects by encryption type" section shows 0 objects for customer-managed, SSE-KMS managed, and SSE-S3 managed encryption, and 6 objects for no encryption. The "Public access" section shows "有効なアクセス許可" (Valid access permissions) as "Not public". The "Replication" section shows "レプリケート済み" (Replicated) as "いいえ" (No) for both internal and external replication. At the bottom of the sidebar, there are settings for "アカウントレベル" (Account level) and "ACLによるパブリックアクセスをプロ" (Public access by ACL) which is set to "オン" (On).

| バケット   | アカウント | サイズ      | オブジェ... |
|--|-------|----------|---------|
| <input type="checkbox"/> aws-                              |       | 456.8 MB | 36      |
| <input type="checkbox"/> aws-                              |       | 22.9 MB  | 266     |
| <input type="checkbox"/> aws-                              |       | 1.1 MB   | 6       |
| <input type="checkbox"/> aws-                              |       | 1.2 MB   | 1       |
| <input type="checkbox"/> confi                             |       | 56.4 MB  | 2.9 k   |
| <input type="checkbox"/> labda                             |       | 1.5 GB   | 623.8 k |
| <input type="checkbox"/> my-s                              |       | 2.6 GB   | 335.8 k |
| <input type="checkbox"/> satos                             |       | 2.3 MB   | 22      |
| <input type="checkbox"/> satos                             |       | 0        | 0       |
| <input type="checkbox"/> satst                             |       | 0        | 0       |
| <input type="checkbox"/> satst                             |       | 0        | 0       |
| <input type="checkbox"/> satst                             |       | 110.0 MB | 112.3 k |
| <input type="checkbox"/> satstak-internal-s3bucket         |       | 2.0 GB   | 813.4 k |
| <input checked="" type="checkbox"/> satstak-macie-test     |       | 1.2 KB   | 6       |
| <input type="checkbox"/> satstak-sharing                   |       | 9.8 MB   | 6       |
| <input checked="" type="checkbox"/> satstak-trail-test0529 |       | 289.2 MB | 119.8 k |

**satstak-macie-test**

概要

|             |                                 |
|-------------|---------------------------------|
| アカウント ID    |                                 |
| リージョン       | アジアパシフィック (東京) ap-northeast-1   |
| ARN         | arn:aws:s3:::satstak-macie-test |
| 名前          | satstak-macie-test              |
| 作成済み        | 2020年5月25日, 12:35:29 (2 か月 前)   |
| オブジェクト      | 6                               |
| 分類可能なオブジェクト | 6                               |
| 共有アクセス      | Not shared                      |
| サイズ         | 1 KB                            |
| 圧縮サイズ       | 0 Bytes                         |
| バージョンング     | 無効                              |
| 最終分析日       | 2020年8月6日, 11:10:04 (4 時間 前)    |

Number of objects by encryption type

|               |   |
|---------------|---|
| カスタマーマネージド    | 0 |
| SSE-KMS マネージド | 6 |
| SSE-S3 マネージド  | 0 |
| 暗号化なし         | 0 |

Public access

|           |            |
|-----------|------------|
| 有効なアクセス許可 | Not public |
|-----------|------------|

Replication

|               |     |
|---------------|-----|
| レプリケート済み      | いいえ |
| 外部でレプリケート済み   | いいえ |
| レプリケーションアカウント | なし  |

アカウントレベル 設定

|                    |    |
|--------------------|----|
| ACLによるパブリックアクセスをプロ | オン |
|--------------------|----|

## 2. 機微情報の検出・評価



# Macie のスキャンはジョブを作成して定義する



Amazon Macie

## ジョブ定義

ジョブ名

対象 S3 バケット（複数指定可能）

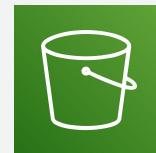
機密データ 検出オプション

実行スケジュール

サンプリング深度

オブジェクト条件

カスタムデータ識別子オプション



Amazon S3



Bucket-A



Bucket-B



Bucket-C

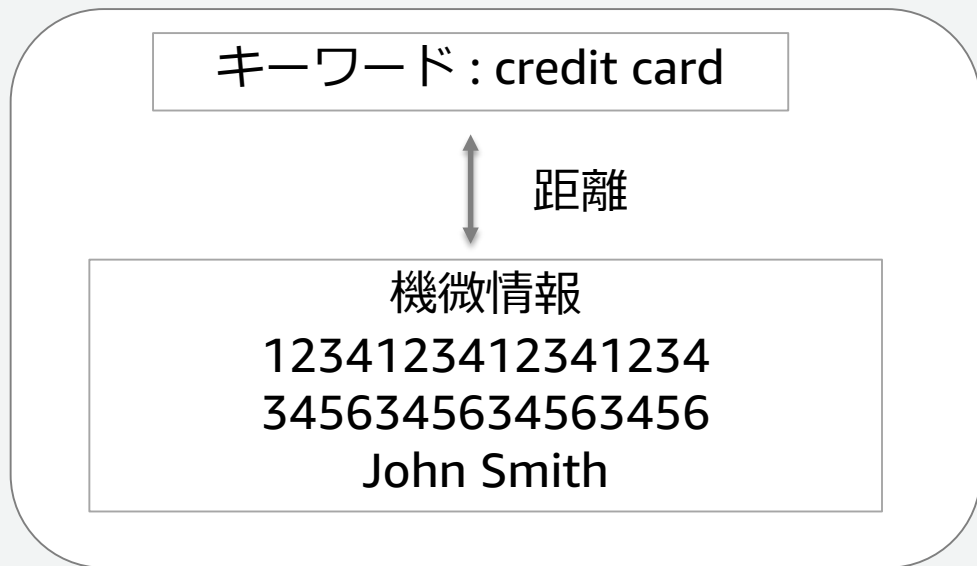
# 参考 : S3 バケット以外のデータをスキャンしたい場合は？

- Amazon Macie は現在 S3 バケットのスキャンのみをサポートしている
- ワークアラウンド
- Amazon RDS を Parquet 形式で S3 バケットにエクスポートして、それをスキャンする
- 他の AWS サービスやデータストアで CSV でエクスポート出来る場合には、それを S3 バケットの上にエクスポートしてスキャンする

# Macie の機微情報検出方法

- Macie は機械学習とパターンマッチングを組み合わせることで機微情報を検出する

ファイル/オブジェクト



1

キーワードが存在するか  
チェック  
※ キーワード不要なケース  
あり

2

キーワードから近い位置  
(距離が短い) に機微情  
報があるかをチェック

3

機微情報に該当する文字  
列があるかをチェック

# スキャンスコープの設定 1

- ワンタイムの実行とスケジュール実行の選択が可能
- スケジュールされたジョブでは、どこまでスキャンをしたかを保存しておくことで、差分スキャンを行わせることが可能

ステップ 1  
S3 バケットを選択

ステップ 2  
S3 バケットを確認

ステップ 3  
スコープ

ステップ 4  
カスタムデータ識別子

ステップ 5  
名前と説明

ステップ 6  
確認して作成

## スコープ Info

これらの設定を使用して、ジョブを実行する頻度を指定します。また、ジョブの分析の深さと範囲を指定することもできます。

### 機密データ検出オプション

スケジュールされたジョブ  
スケジュールされた頻度でオブジェクトを分類

更新の頻度  
毎日

ワンタイムジョブ  
既存のオブジェクトを 1 回だけ分類

既存のオブジェクトを含める  
新しいオブジェクトと既存のオブジェクトを分類するには、このオプションを選択します。新しいオブジェクトのみを分類する場合は、このオプションをオフにします。

サンプリング深度  %  
深度パーセンテージに基づいてオブジェクトの

▶ その他の設定

前へ 次へ

スケジュールされたジョブでは、スキャンしたオブジェクトをマークして、次のスケジュールではマーク済みのオブジェクトを除外可能

# 設定可能なスケジュール

- 毎日、毎週、毎月を選択可能
- ジョブが実行される時間帯や日付を指定することは出来ない

# サンプリング深度とは？

- ジョブの中でサンプリング深度を指定（デフォルトは 100）
- 100（=デフォルト）を指定すると、全てのオブジェクトをスキャンする
- 100 未満を指定すると、ランダムに指定された比率のオブジェクトをスキャンする。100オブジェクト存在した場合に、90を指定すると90オブジェクトをランダムに選択してスキャンする

# スキャンスコープの設定 2

- スキャン対象にする条件、除外する条件を設定することで、スキャン対象を絞り込む設定が可能
- 条件には、タグ、最終改訂日、ファイル拡張子、オブジェクトサイズを利用可能
- 除外条件は対象にする条件を上書きする（両方マッチした場合は除外される）

▼ その他の設定

左側のリストから少なくとも1つの条件を選択し、「含める」または「除外」をクリックします。条件を追加しない場合、すべてのオブジェクトが分類されません。

オブジェクト条件

ファイル拡張子 ▼

拡張子を区切るには、「,」を使用します。たとえば、「pdf, zip, doc」などです。

含める

ファイル拡張子 : json 削除

除外

除外条件が追加されていません

含める 除外

JSON ファイルのみ  
対象と設定した例

# Macie のデータ識別子

## AWS マネージド識別子

- AWS がマネージして提供する機微情報検出用の定義
- お客様によるカスタマイズは出来ない
- Macie のジョブを実行すると自動で検出が行われる

## カスタマーマネージド識別子

- お客様が検出定義を作成して独自の検出を行うことが可能
- キーワードと正規表現のパターンマッチング定義を作成する必要がある
- ジョブ単位に指定を行う必要がある（作成した識別子定義は再利用可能）



# AWS マネージド識別子の対象情報 1 (抜粋)

| カテゴリ            | データタイプ<br>(抜粋) | キーワード  | 補足情報              | 対応している国と<br>リージョン                             |
|-----------------|----------------|--|-------------------|---|
| 個人特定情報<br>(PII) | 生年月日           | Bday, b-day, birth date,<br>birthday, date of birth, dob   |                   | 全て  |
|                 | 姓名             | 必要なし   | ラテン言語の文字<br>セットのみ | 全て  |
|                 | 住所             | 必要なし   | ラテン言語の文字<br>セットのみ | 全て  |
|                 | パスポート番号        | <a href="https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html#managed-data-identifiers-pii-passport-keywords">https://docs.aws.amazon.com/macie/latest/<br/>user/managed-data-<br/>identifiers.html#managed-data-identifiers-<br/>pii-passport-keywords</a> |                   | カナダ、フランス、<br>ドイツ、イタリア、<br>スペイン、UK、US          |
|                 | 電話番号           | cel, cell, celular, contact, fone,<br>mobile, móvel, número<br>residencial, numero<br>residencial, phone, tel,<br>telephone, telephone number  |                   | ブラジル、カナダ、<br>フランス、ドイツ、<br>イタリア、スペイン、<br>UK、US |

参考: <https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html>

# AWS マネージド識別子の対象情報 2 (抜粋)

| カテゴリ   | データタイプ<br>(抜粋)      | キーワード   | 補足情報   | 対応している国と<br>リージョン            |
|--------|---------------------|---|--|------------------------------|
| 金融関連情報 | 銀行口座番号              | <a href="https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html#managed-data-identifiers-financial-bankacct-keywords">https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html#managed-data-identifiers-financial-bankacct-keywords</a> |  | カナダ、フランス、ドイツ、イタリア、スペイン、UK、米国 |
|        | クレジットカード失効日         | expiration, expiry  | サポートされるフォーマットは MM/YY もしくは YY/MM                                | 全て                           |
|        | クレジットカードの磁気ストライプデータ | ard data, iso7813, mag, magstripe, stripe, swipe  | Tracks 1 と 2を含む  | 全て                           |
|        | クレジットカード番号          | なし  | キーワードは不要だが、ルーンアルゴリズムによるチェックディジットを含む <b>13桁から19桁</b> のフォーマットが必要 | 全て                           |

参考: <https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html>

# AWS マネージド識別子の対象情報 3 (抜粋)

| カテゴリ | データタイプ<br>(抜粋) | キーワード  | 補足情報 | 対応している国と<br>リージョン |
|------|----------------|--|------|-------------------|
| 認証情報 | AWS シークレットキー   | aws_secret_access_key,<br>credentials, secret access key,<br>secret key, set-awscredential |      | 全て                |
|      | OpenSSH 秘密鍵    | なし   |      | 全て                |
|      | PGP 秘密鍵        | なし   |      | 全て                |
|      | PKCS 秘密鍵       | なし   |      | 全て                |
|      | PuTTY 秘密鍵      | なし   |      | 全て                |

参考: <https://docs.aws.amazon.com/macie/latest/user/managed-data-identifiers.html>

# 日本語対応

- サポートされている文字コード
  - UTF-8 のみがサポートされている
  - 他の文字コード (Shift\_JIS、EUC) の文字が入ったファイルはスキャン対象外となる (スキップされる)
- AWS マネージドの識別子では、日本語のキーワードが登録されていない
  - 例 Credit Card というキーワードがあると検出するが “クレジットカード” というキーワードは検出しない
- カスタマー マネージドの識別子では、日本語のキーワードを指定可能
  - マネージメントコンソールで日本語のキーワードを指定した場合、UTF-8として解釈される、正規表現で日本語の文字とのマッチングを指定した場合もUTF-8の場合はマッチングが行われる

# カスタムデータ識別子

- Macie は PCRE (Perl Compatible Regular Expressions) を利用した正規表現によるパターンマッチングを利用してお客様のユースケースに応じたカスタム検出をサポート
- カスタムデータ識別子で指定可能なパターン
  - キーワード
  - 除外キーワード
  - 最大一致距離
  - 正規表現パターン

参照 <https://docs.aws.amazon.com/macie/latest/user/custom-data-identifiers.html>

# カスタムデータ識別子の定義

Macie > カスタムデータ識別子 > CreditCardNumber

**CreditCardNumber** Info

Id  
a41e6038-4a75-4f0c-b03d-a21dce8e3fff

作成日  
2020年6月1日, 14:47:06 (2 か月 前)

説明  
The format for credit card number

**正規表現**  
一致するパターンを定義する正規表現 (regex) を入力します。  
^(?:[0-9]{12}|(?:[0-9]{3})?5[1-5][0-9]{14}|6011[0-9]{12}|3(?:0[0-5]||[68][0-9])[0-9]{11})3[47][0-9]{13}|(?:2131|1800|35[0-9]{5})[0-9]{11}

**キーワード**

**単語を無視**

最大一致距離  
50

**評価**

サンプルデータ  
3540120000000000

送信

結果  
2 1 個の一致

正規表現によるマッチングが正しく動作するか実際のデータを入力して確認可能

正規表現パターン

キーワード

無視キーワード

キーワードと機微情報の距離 (単位: 文字数)  
最大 300まで指定可能

# 正規表現による機微情報のマッチング例

## クレジットカード番号

VISA の例、4で始まる16桁の数字

```
^4[0-9]{12}(?:[0-9]{3})?$
```

## IP アドレス

例 192.0.2.5

```
^(?:(?:25[0-5]|2[0-4][0-9]|[01]?[0-9][0-9]?)¥.){3}(?:25[0-5]|2[0-4][0-9]|[01]?[0-9][0-9]?)$
```

※ Macie がサポートしない正規表現要素

- 前方参照
- キャプチャグループ
- 条件つきパターン
- 埋め込みコード
- /i, /m, /x のようなグローバルフラグ
- 再帰的パターン
- ?, !, ?<=, ?<! などの肯定先読み、否定後読み

# 暗号化された S3 バケットのスキャン

- SSE-S3, AWS-KMS, SSE-KMS による暗号化が行われている場合、Macie はバケットをスキャンすることが可能
  - SSE-S3, AWS-KMS の場合は追加設定不要でスキャン可能
  - SSE-KMS の場合には、Macie のサービスリンクロール `AWSServiceRoleForAmazonMacie` に CMK へのアクセス権限の付与が必要
- クライアントサイド暗号化、カスタマー提供型のサーバーサイド暗号化 (SSE-C) の場合には、復号する手段が無いため、Macie はスキャンが出来ない
  - バケットの使用状況の可視化は S3 のメタデータを使っているため、上記のタイプの暗号化がされていても実行可能



# サポートされるファイルタイプ

| ファイル/<br>ストレージタイプ | 説明   | ファイル拡張子                                       |
|-------------------|--|---|
| ビッグデータ            | Apache Avro オブジェクトコンテナ、<br>Apache Parquet ファイル                             | N/A   |
| 圧縮・アーカイブ          | GNU Zip 圧縮アーカイブ、Tar アー<br>カイブ、ZIP 圧縮アーカイブ                                  | .gz, .gzip, .tar, .zip                        |
| 文書                | Adobe PDF ファイル、Microsoft エ<br>クセルワークブック、Microsoft<br>Word 文書               | .doc, .docx, .pdf, .xls, .xlsx                |
| テキスト              | CSV ファイル、HTMLファイル、<br>JSONファイル、プレーンテキスト<br>文書、タブ区切りファイル (TSV) 、<br>XMLファイル | .csv, .htm, .html, .json, .tsv,<br>.txt, .xml |

※ Macie は画像、オーディオ、ビデオなどのメディアファイルは分析しない、また、文字コードやサイズの制限があるので注意する

参考 : [https://docs.aws.amazon.com/ja\\_jp/macie/latest/user/discovery-supported-formats.html](https://docs.aws.amazon.com/ja_jp/macie/latest/user/discovery-supported-formats.html)

### 3. 評価・検出結果の参照、 他のサービスとの連携

# マネジメントコンソール上での検出結果の確認

検出ジョブの実行が完了すると、マネジメントコンソール上で検出結果の確認が出来る

Macie > 検出結果 表示中: 10 / 10 1 2 7

検出結果 (10)  アクション

このテーブルには、組織についての検出結果が一覧表示されます。検出結果を選択して詳細を表示します。また、特定のフィールドとフィールド値に基づいて検出結果をフィルタリング、グループ化、ソートすることもできます。

保存済みフィルタ / 自動アーカイブ  保存済みのフィルタがありません

現在  フィルタの追加

| 重...   | 検出結...           | 影響を受けるリソース                                | 更新日     | カ... |
|--------|------------------|---|---------|------|
| High   | SensitiveData... | satstak-macie-test/macie_test_data4.txt   | 3 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data3.txt   | 3 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data4.txt   | 3 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data3.txt   | 3 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data3.txt   | 3 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data3.txt   | 4 日間 前  | 1    |
| High   | SensitiveData... | satstak-macie-test/macie_test_data3.txt   | 4 日間 前  | 1    |
| Medium | SensitiveData... | satstak-macie-test/macie_test_data2.txt   | 4 日間 前  | 1    |
| Low    | SensitiveData... | satstak-internal-s3...giwi7ALxHh7.json.gz | 10 日間 前 | 1    |
| Medium | [サンプル] Po...     | create-sample-finding-bucket              | 21 日間 前 | 2    |

検出結果の一覧、重大度は自動的に割り当てられる

**SensitiveData:S3Object/Financial** 検出結果 ID: 20ef8e75b26d60ae05e217fdb18f7259

**High** The object contains financial information such as credit card numbers or bank account numbers. [Learn More](#)

概要

|          |   |
|----------|---|
| 重大度      | High                                    |
| リージョン    | ap-northeast-1                          |
| アカウント ID | [REDACTED]                              |
| 資源       | satstak-macie-test/macie_test_data3.txt |
| 作成日      | 06-02-2020 18:29:01 (3 日間 前)            |
| 更新日      | 06-02-2020 18:29:01 (3 日間 前)            |

結果

|        |                                  |
|--------|----------------------------------|
| ジョブ ID | 730f40340e2573e9b621c2dbccc6e93f |
|--------|----------------------------------|

詳細

|          |   |
|----------|---|
| ステータス    | COMPLETE  |
| 分類されたサイズ | 309 Bytes   |
| MIME タイプ | text/plain  |
| 詳細な結果の場所 | s3://[export-config-not-set]/AWSLogs/[REDACTED]/... |

財務情報

|                    |   |
|--------------------|---|
| Credit card number | 2 |
|--------------------|---|

影響を受けるリソース (S3 バケット)

|           |                               |
|-----------|-------------------------------|
| バケット名     | sa                            |
| パブリックアクセス | N                             |
| 暗号化タイプ    | aws:kms                       |
| 作成日       | 05-25-2020 12:35:29 (11 日間 前) |
| オーナー      | satstak                       |

クレジットカード番号が2件含まれていた

- ファイル単位に検出結果が出力される
- 何も検出されなかった場合には出力されない
- 1つのファイルに複数の機微情報が入っていた場合には1つの結果が検出結果が出力され、その中にいくつかのデータがあったかが表示される
- 何行目に含まれていたという情報は検出結果に含まれない

# 検出結果の非表示化

検出結果をフィルター表示する画面で、指定した条件の Findings を自動アーカイブする Suppression Rule を作成して適用することが可能。不要な表示を減らすことで、管理を容易にする。

検出結果 (0) [Info](#) 🔄 アクション ▼

このテーブルには、組織についての検出結果が一覧表示されます。検出結果を選択して詳細を表示します。また、特定のフィールドとフィールド値に基づいて検出結果をフィルタリング、グループ化、ソートすることもできます。

保存済みフィルタ / 自動アーカイブ 保存済みのフィルタがありません ▼

現在 ▼  ⊗ Add filter ✕

**Suppression rule**  
To automatically suppress findings that meet certain criteria, enter the criteria in the filter bar, and then save the criteria as a suppression rule.

名前

説明 - optional

After you create a suppression rule, Macie continues to generate findings that meet the criteria. However, Macie archives the findings automatically and stops publishing the findings as Amazon CloudWatch Events.

キャンセル Save

重大度 ▼  検出結果タイプ ▼  影響を受けるリソース ▼  更新日 ▼

フィルタ設定により、すべての検出結果が非表示になっています。  
選択されたフィルタにより、すべての検出結果が非表示になっています。検出結果を表示するには、フィルタを変更してください。

## 適用出来る条件（抜粋）

- アカウントID
- カテゴリ
- 検出結果タイプ
- 重大度
- S3 オブジェクトタグキー、タグ値
- S3 バケットタグキー、タグ値
- S3 バケット名

# アクションの実行

- 検出結果を選択して、アクションメニューからJSON へのエクスポートが可能
- ブラウザ経由でローカルにダウンロードされる
- 同じ内容を S3 に自動保存する設定をすることが可能
  - Macie の検出結果は、30日間のみ保管されているので、長期間保管したい場合には、S3 に保存することを推奨する

```
検出結果 JSON
```

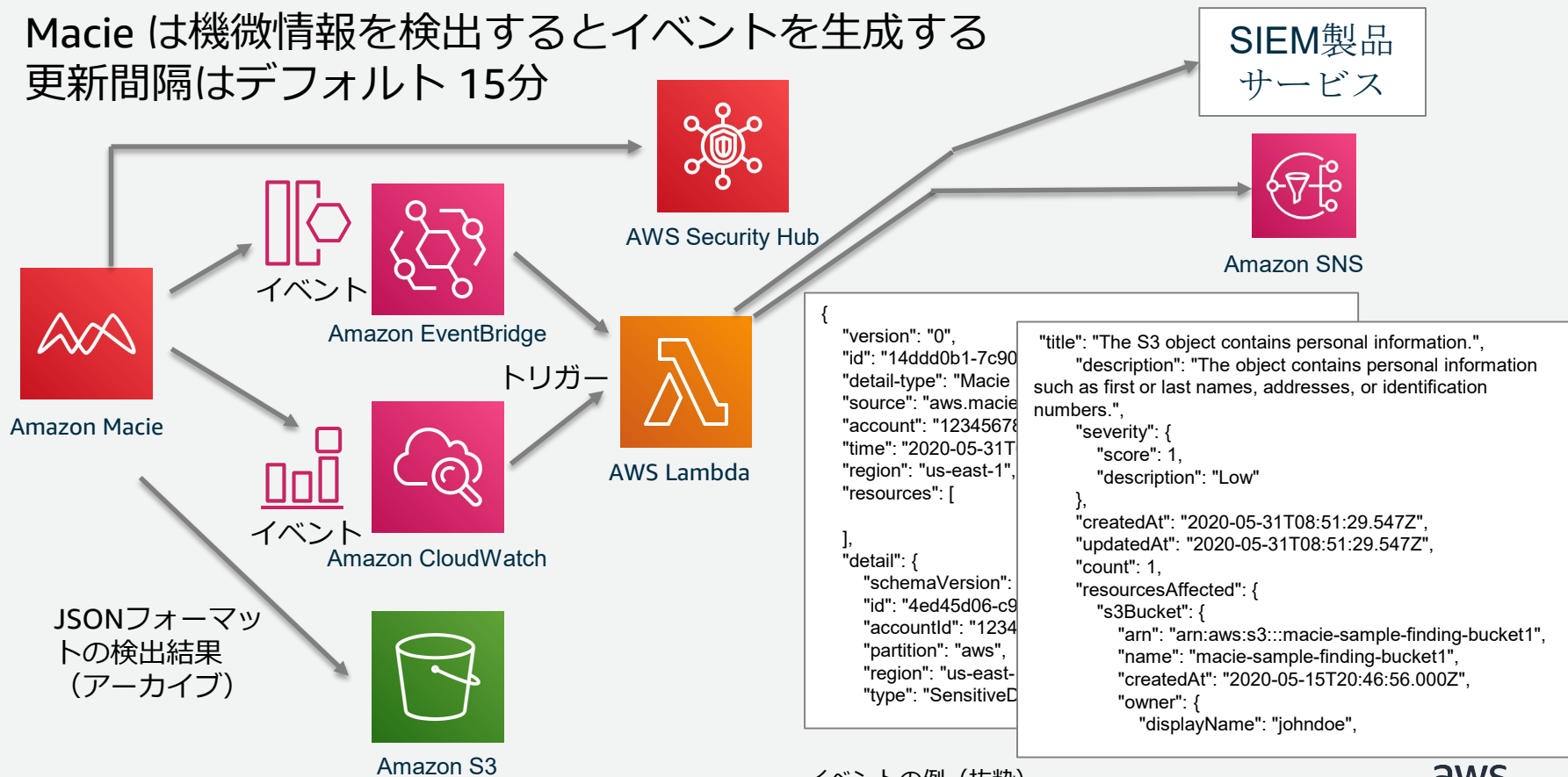
読み取り専用 ⓘ

```
1 - [
2 - {
3   "accountId": "██████████",
4   "archived": false,
5   "category": "CLASSIFICATION",
6 -   "classificationDetails": {
7     "detailedResultsLocation": "s3://[export-config-not-set]/AWSLogs/██████████
      /Macie/ap-northeast-1/f69f6f69bc5b206cc7441f05895b0ac5/12a2c1d8-73f6-3923
      -b6fc-9d157b6ef835/",
8     "jobArn": "arn:aws:macie2:ap-northeast-1:164348464951:classification-job
      /f69f6f69bc5b206cc7441f05895b0ac5",
9     "jobId": "f69f6f69bc5b206cc7441f05895b0ac5",
10 -   "result": {
11 -     "customDataIdentifiers": {
12 -       "detections": [],
13 -       "totalCount": 0
14 -     },
15 -     "mimeType": "application/gzip",
16 -     "sensitiveData": [
17 -       {
18 -         "category": "FINANCIAL_INFORMATION",
19 -         "detections": [
20 -           {
21 -             "count": 1,
22 -             "type": "CREDIT_CARD_NUMBER"
23 -           }
24 -         ]
25 -       }
26 -     ]
27 -   }
28 - }
29 - ]
```

キャンセル [ダウンロード](#)

# イベントの生成と AWS サービスとの連携

Macie は機微情報を検出するとイベントを生成する  
更新間隔はデフォルト 15分

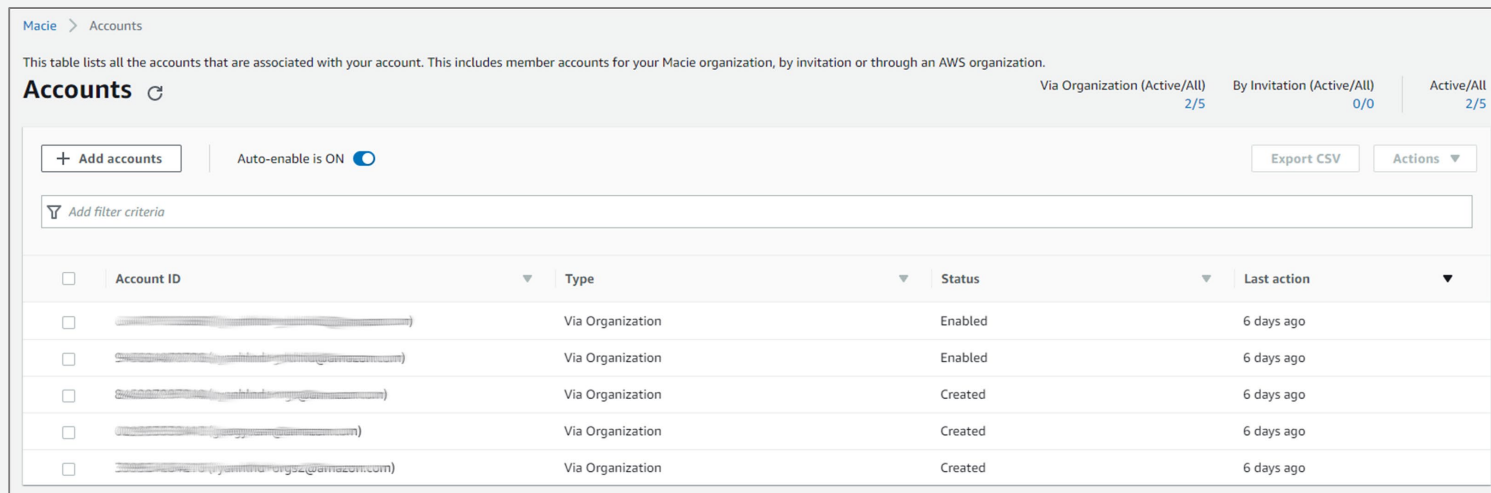


イベントの例 (抜粋)

## 4. 管理、運用、制約

# マルチアカウント環境における Macie の管理

- マスター/メンバー構成をとることが可能
- マスターアカウントの Macie でジョブを作成して、メンバーアカウントで実行出来る
- 手動によるメンバーアカウント追加の場合 1000 アカウントまでサポート
- AWS Organizations 構成の場合は 5000 アカウントまでサポート



Macie > Accounts

This table lists all the accounts that are associated with your account. This includes member accounts for your Macie organization, by invitation or through an AWS organization.

### Accounts

Via Organization (Active/All) 2/5 | By Invitation (Active/All) 0/0 | Active/All 2/5

+ Add accounts | Auto-enable is ON  | Export CSV | Actions

Add filter criteria

| <input type="checkbox"/> | Account ID | Type             | Status  | Last action |
|--------------------------|------------|------------------|---------|-------------|
| <input type="checkbox"/> | [REDACTED] | Via Organization | Enabled | 6 days ago  |
| <input type="checkbox"/> | [REDACTED] | Via Organization | Enabled | 6 days ago  |
| <input type="checkbox"/> | [REDACTED] | Via Organization | Created | 6 days ago  |
| <input type="checkbox"/> | [REDACTED] | Via Organization | Created | 6 days ago  |
| <input type="checkbox"/> | [REDACTED] | Via Organization | Created | 6 days ago  |



# Macie API コールのロギング

- CloudTrail への API コール記録をサポートしている

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "123456789012",
    "arn": "arn:aws:iam::123456789012:user/Mary",
    "accountId": "123456789012",
    "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
    "userName": "Mary",
  },
  "eventTime": "2020-05-22T16:09:56Z",
  "eventSource": "macie2.amazonaws.com",
  "eventName": "ListFindings",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "198.51.100.1",
  "userAgent": "console.amazonaws.com",
  "requestParameters": {
    "sortCriteria": {
      "attributeName": "updatedAt",
      "orderBy": "DESC"
    }
  },
}
```

```
"findingCriteria": {
  "criteria": {
    "archived": {
      "eq": [
        "false"
      ]
    },
    "category": {
      "eq": [
        "POLICY"
      ]
    }
  }
},
"maxResults": 10
},
"responseElements": null,
"requestID": "d58af6be-1115-4a41-91f8-ace03EXAMPLE",
"eventID": "ad97fac5-f7cf-4ff9-9cf2-d0676EXAMPLE",
"eventType": "AwsApiCall",
"recipientAccountId": "123456789012"
}
```

# API, AWSCLI からの呼び出し

新しい Macie は サービスとして macie2 を指定して API や AWS CLI から呼び出しが可能

※ 東京リージョンで利用出来なかった 前のバージョンの Macie (v1) があるため、2 がついている

```
$ aws macie2 list-findings
{
  "findingIds": [
    "495eb82347841ba75f56f82cc35ba5bd",
    "c8c99a7b6617b43c1e7ad459555120b0",
    "b0e0cd61f868a825307b7f69c04e8e56",
    "f939afe6c315636b44d6cd40620e777a",
    "7e7c1318d2815e122ebd92243460c264",
    "ce64ea47a88a25558e7b6a39e7271f14",
    "58f8487fd7091a9949cc715804715d14"
  ]
}
```

```
$ aws macie2 get-findings --finding-ids 495eb82347841ba75f56f82cc35ba5bd
{
  "findings": [
    {
      "accountId": "000000000000",
      "archived": false,
      "category": "CLASSIFICATION",
      "classificationDetails": {
        "detailedResultsLocation": "s3://[export-config-not-
set]/AWSLogs/164348464951/Macie/ap-northeast-
1/f69f6f69bc5b206cc7441f05895b0ac5/12a2c1d8-73f6-3923-b6fc-9d157b6ef835/",
        "jobArn": "arn:aws:macie2:ap-northeast-1:164348464951:classification-
job/f69f6f69bc5b206cc7441f05895b0ac5",
        "jobId": "f69f6f69bc5b206cc7441f05895b0ac5",
        "result": {
          以下省略
        }
      }
    }
  ]
}
```

参照: <https://docs.aws.amazon.com/cli/latest/reference/macie2/index.html>

# Macie のセキュリティ

- Macie は Amazon S3 のデータをスキャンするが、データそのものや、スキャン結果が勝手に利用されることはないか？
  - Macie によるスキャンデータは解析のためにメモリー上で一時的に保持されるが、他のストレージにコピーされることはない
  - スキャン以外の目的で Macie のサービスは、お客様の S3 のデータにアクセスしたり、スキャン結果を利用することはない
  - Macie によるスキャン、分類結果は、暗号化されたストレージに保管される
- Amazon S3 と Macie の間の通信は暗号化される

参考 : <https://docs.aws.amazon.com/macie/latest/user/data-protection.html>

# サポートされるリージョン (2020年8月現在)

- Macie はリージョン別のサービスで、リージョンごとに有効化する必要がある
- Macie を利用可能なリージョン
  - 米国 (GovCloudを除く)、サンパウロ、東京、シンガポール、シドニー、ソウル、ムンバイ、香港、フランクフルト、アイルランド、ロンドン、パリ、ストックホルム

サポートされるリージョンの情報

<https://aws.amazon.com/jp/about-aws/global-infrastructure/regional-product-services/>

# 制約、クォータ

## 制約

- メンバーアカウントの数 : 手動の場合 1,000 アカウント、 AWS Organizations 経由の場合 5,000 アカウント
- 検出ジョブ単位の Findings の数 100,000 + 検出対象のオブジェクト数の 5% が限度
- 機微情報の検出と分類
  - 月あたりの検出と分類の上限 5TB、 サービスクォータコンソールから25TB までの増加をリクエスト可能。さらに増やす場合にはサポートにご相談ください
  - データ検出結果中の検出口ケーション : 検出タイプごとに 1,000
  - データ検出ジョブに設定出来るカスタムデータ識別子の数 : 30
  - 圧縮されたりアーカイブされているファイルの解凍の上限 :  
ネストされたアーカイブの深さの制限 : 10 、 解凍するファイル数 1,000,000
  - ファイルサイズの制限 : PDF 1024MB, テキストファイル 20GB, Microsoft Excel /Word 512MB (一部を抜粋、詳細は下記URL参照)

<https://docs.aws.amazon.com/macie/latest/user/macie-quotas.html>

# 利用料金

- 課金対象
  - S3 バケットの数 (月ごと)
  - 機微情報検出対象となったデータ処理量 (月ごと)
- 東京リージョンの例
  - S3 バケットあたり 0.10 USD
  - データ処理量
    - 50,000GB /月までGBあたり 1.25 USD
    - 450,000GB /月までGBあたり 0.63 USD
    - 500,000DB /月を超えた場合、GBあたり0.31USD
- Macie が S3 にアクセスする際の、GET および LIST について S3側の料金もかかる
- 無料試用期間 30日
  - バケットの可視化のみが対象で、スキャンに関しては下記のルールが適用される
  - スキャンに関して、1ヶ月 1GBまで無料 (30日間限定ではない)

# まとめ

Amazon Macie を利用することで

S3 バケット上の大量の機微情報の検出を効率的に実行可能

容易な設定と管理が実現可能

AWS マネージド識別子と、カスタム識別子の両方を  
活用することで効率よく検出が可能

スケジュールして定期的にスキャンを行うことで、S3 バケット上に  
意図しない機微情報が含まれていることがないことを可視化可能

# 参考資料

- Amazon Macie ドキュメント

[https://docs.aws.amazon.com/ja\\_jp/maciek/index.html](https://docs.aws.amazon.com/ja_jp/maciek/index.html)

※ 2020年8月現在 英語のみの提供



# AWS の日本語資料の場所「AWS 資料」で検索



The screenshot shows the AWS Japanese website header with the logo, navigation links for '日本語' and 'アカウント', and a 'サインイン' button. The main content area features the title 'AWS クラウドサービス活用資料集トップ' and a paragraph of introductory text. Below the text are four navigation buttons: 'AWS Webinar お申込', 'AWS 初心者向け', '業種・ソリューション別資料', and 'サービス別資料'.

aws

日本担当チームへお問い合わせ サポート 日本語 アカウント

コンソールにサインイン

製品 ソリューション 料金 ドキュメント 学習 パートナー AWS Marketplace その他

## AWS クラウドサービス活用資料集トップ

アマゾン ウェブ サービス (AWS) は安全なクラウドサービスプラットフォームで、ビジネスのスケールと成長をサポートする処理能力、データベースストレージ、およびその他多種多様な機能を提供します。お客様は必要なサービスを選択し、必要な分だけご利用いただけます。それらを活用するために役立つ日本語資料、動画コンテンツを多数ご提供しております。(本サイトは主に、AWS Webinar で使用した資料およびオンデマンドセミナー情報を掲載しています。)

AWS Webinar お申込 »

AWS 初心者向け »

業種・ソリューション別資料 »

サービス別資料 »

<https://amzn.to/JPArchive>

# Q&A

お答えできなかったご質問については

AWS Japan Blog 「<https://aws.amazon.com/jp/blogs/news/>」にて

後日掲載します。

# AWS Well-Architected 個別技術相談会

毎週“W-A個別技術相談会”を実施中

- AWSのソリューションアーキテクト(SA)に  
対策などを相談することも可能

• 申込みはイベント告知サイトから

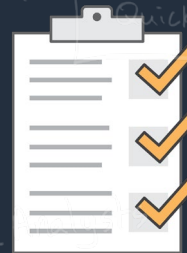
(<https://aws.amazon.com/jp/about-aws/events/>)

AWS イベント

で[検索]



AWS Well-Architected



# ご視聴ありがとうございました

AWS 公式 Webinar

<https://amzn.to/JPWebinar>



過去資料

<https://amzn.to/JPArchive>

