National College of Ireland

# Configuration Manual

MSc Research Project
Cyber Security

Cormac Frawley
Student ID: X21166277

School of Computing
National College of Ireland

Supervisor: Raza Ul Mustafa

**National College of Ireland**

**MSc Project Submission Sheet**

**School of Computing**

| | |
|---|---|
| **Student Name:** | Cormac Frawley |
| **Student ID:** | X21166277 |
| **Programme:** | MSc Cyber Security – Part Time **Year:** Jan 2022 |
| **Module:** | MSc Research Project |
| **Lecturer:** | Raza Ul Mustafa |
| **Submission Due Date:** | 14th December 2023 |
| **Project Title:** | Understanding the Efficacy of Cloud Native Sensitive Data Protection Capabilities |
| **Word Count:** | 2374 **Page Count: 23** |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** .....

**Date:** .....

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies) | □ |
| **Attach a Moodle submission receipt of the online project submission,** to each project (including multiple copies). | □ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | □ |

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Configuration Manual

Cormac Frawley
Student ID: X21166277

# 1 AWS

A Standalone  AWS account was created for the purposes of testing the below configuration. All infrastructure provisioned on the AWS account was through Terraform.

## 1.1 Terraform Code

All Terraform code was written and executed in Visual Studio Code

### 1.1.1 providers.tf

The below code is used to form an authenticated connection to AWS so that the infrastructure and Sensitive Data Protection function can be configured. For AWS the credentials where stored in a directory and accessed once the code was executed. The AWS region is also specified.

```
terraform {
  required_providers {
    aws = {
      source = "hashicorp/aws"
    }
  }
}

# Provider & Credentials
provider "aws" {
  region                  = "eu-west-1"
  shared_credentials_files = ["~/.aws/credentials"]
  profile                 = "cmac_thesis"
}
```

Figure 1 – AWS providers.tf code

### 1.1.2 main.tf

The main.tf file is the main configuration file and the Terraform code executed within it brings up the required AWS cloud environment, storage element (S3 bucket), networking and adds the test files to the S3 bucket.

For AWS the below infrastructure was needed to facilitate the testing:

- VPC (virtual private cloud) is a virtual network within which infrastructure can be instantiated.
- Subnet – A VPC subnet is required to allocate Ipv4 addresses and the availability zone (area within a region where the resources are located) is also defined.
- S3 bucket – a standard S3 bucket is defined.
- S3 object – the S3 object defines a process to upload files from a specified location to the S3 bucket. Each time the bucket is created the object code executes and uploads the correct files.

```
# Create a VPC
resource "aws_vpc" "cmac_dlp_vpc" {
  cidr_block           = "10.10.0.0/16"
```

```
  enable_dns_hostnames = true
  enable_dns_support   = true

  tags = {
    Name = "cmac_vpc"
  }
}

#Create a subnet for the VPC - vpc.id is defined at VPC creation
resource "aws_subnet" "cmac_dlp_vpc_subnet" {
  #vpc_id                  = var.vpc_id
  vpc_id                  = aws_vpc.cmac_dlp_vpc.id
  cidr_block              = "10.10.1.0/24"
  map_public_ip_on_launch = true
  availability_zone       = "eu-west-1a"

  tags = {
    Name = "cmac_vpc_subnet"
  }
}


#Create an S3
resource "aws_s3_bucket" "s3-dlp" {

    bucket = "s3-dlp"

      tags = {
    Name = "S3-dlp-bucket"
  }


}

#Add files to S3
resource "aws_s3_object" "object1" {

    for_each = fileset("uploads/", "*")
    bucket = aws_s3_bucket.s3-dlp.id
    key = each.value
    source = "uploads/${each.value}"
}
```

Figure 2 – AWS providers.tf code

### 1.1.3 dlp.tf

To enable the Sensitive Data Protection capabilities the dlp.tf file enables Macie (Sensitive Data Protection function) and defines the account id for reference in the Macie policy. There are a number of ways of creating the Macie policy, in the code referenced a one-time job was used as this executed immediately which was

useful for testing (AWS, 2023). The code also references a scheduled job type which would allow the policy to trigger on a regular basis and would be used in production code. The policy references the S3 bucket but again this could be configured to reference a number of buckets as required.

Finally Terraform/AWS configuration does not allow configuration of individual sensitive information types (SITs) instead the code uses the default information types defined for Macie. For those SITs outside the default group policies had to be created directly on the portal. This also included the custom information types. The configuration for these will be referenced in the following section.

```
#Turn on Macie
resource "aws_macie2_account" "cmac-macie-dlp" {
  status = "ENABLED"
}

#Returns account id needed for policy
data "aws_caller_identity" "current" {}

#Macie policy
resource "aws_macie2_classification_job" "cmac-macie-job-01" {
  # job_type = "SCHEDULED"
  # name     = "cmac-macie-detection-job"
  # schedule_frequency {
  #   daily_schedule = "true"
  # }

  job_type = "ONE_TIME"
  name     = "cmac-macie-detection-job-01"

  s3_job_definition {
    bucket_definitions {
      account_id = data.aws_caller_identity.current.account_id
      buckets    = [aws_s3_bucket.s3-dlp.id]
    }
  }
  depends_on = [aws_macie2_account.cmac-macie-dlp]
}
```

Figure 3 – AWS dlp.tf code

### 1.1.4 AWS Macie configuration for non default information types

For those SITs not part of the default group the Macie policies had to be configured in the AWS portal. The AWS portal provides a web based GUI to configure all elements of the AWS environment from storage to networking to access management. In this case it was used to add specific SITs to the Macie policy. Below illustrates part of the GUI configuration.

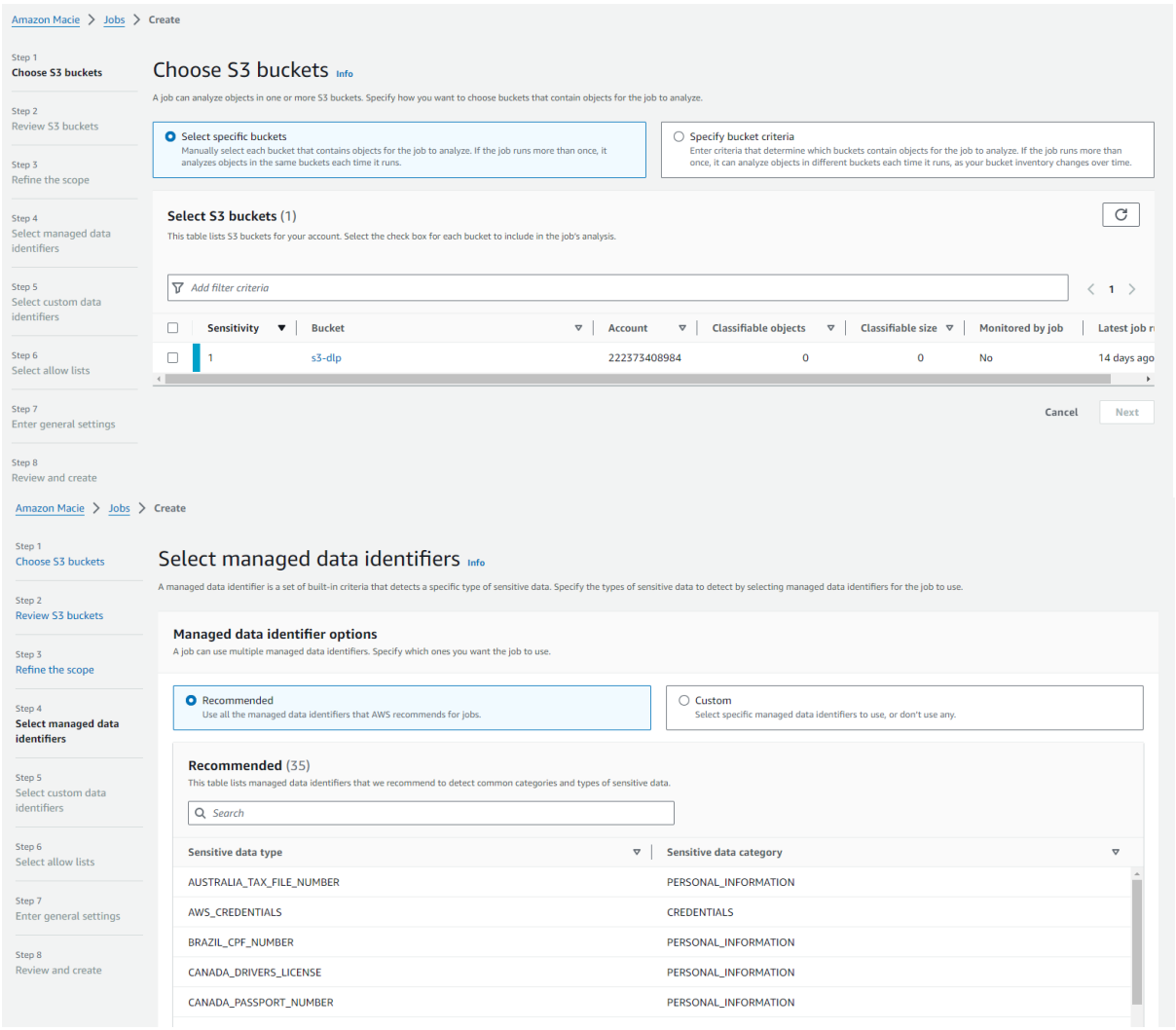


Figure 4 – AWS Macie portal configuration

Similarly the custom information types must be configured in the portal before being added in the Macie policy through the portal. The below configuration shows a custom identifier using regex.

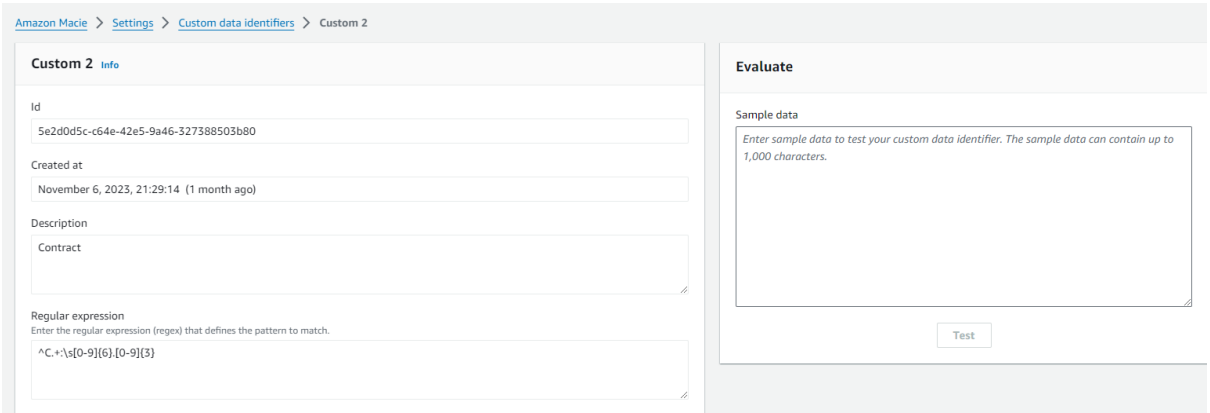


Figure 5 – AWS Macie custom identifier configuration

## 1.2 Macie Sensitive Data Protection Jobs

Once configured the Macie job can be checked in the AWS portal example below.

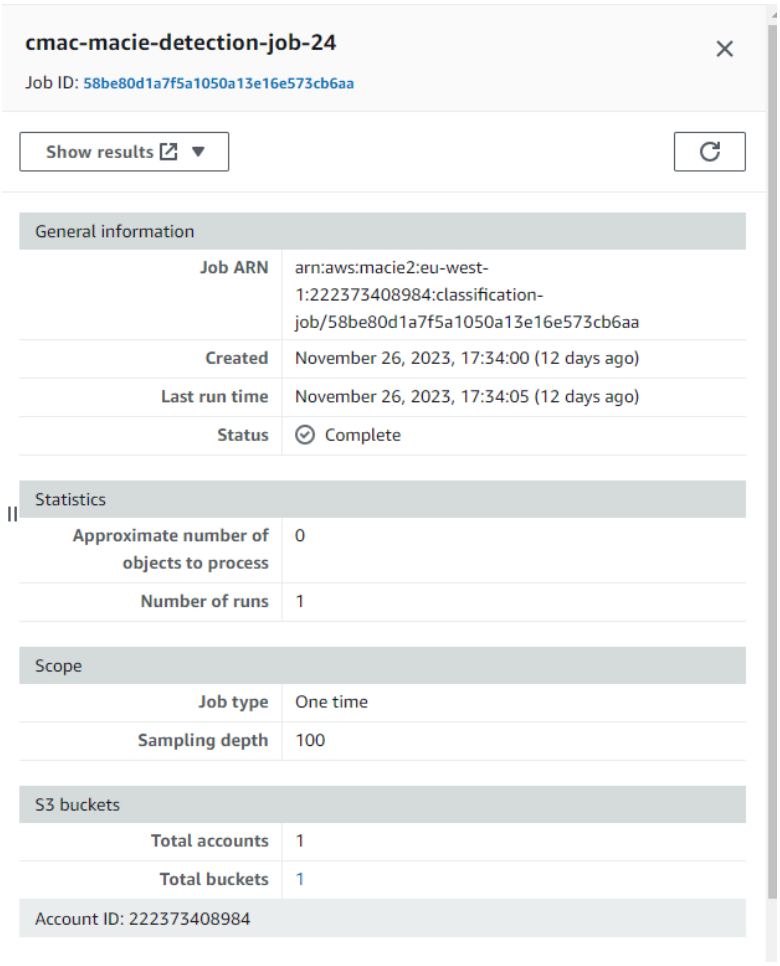
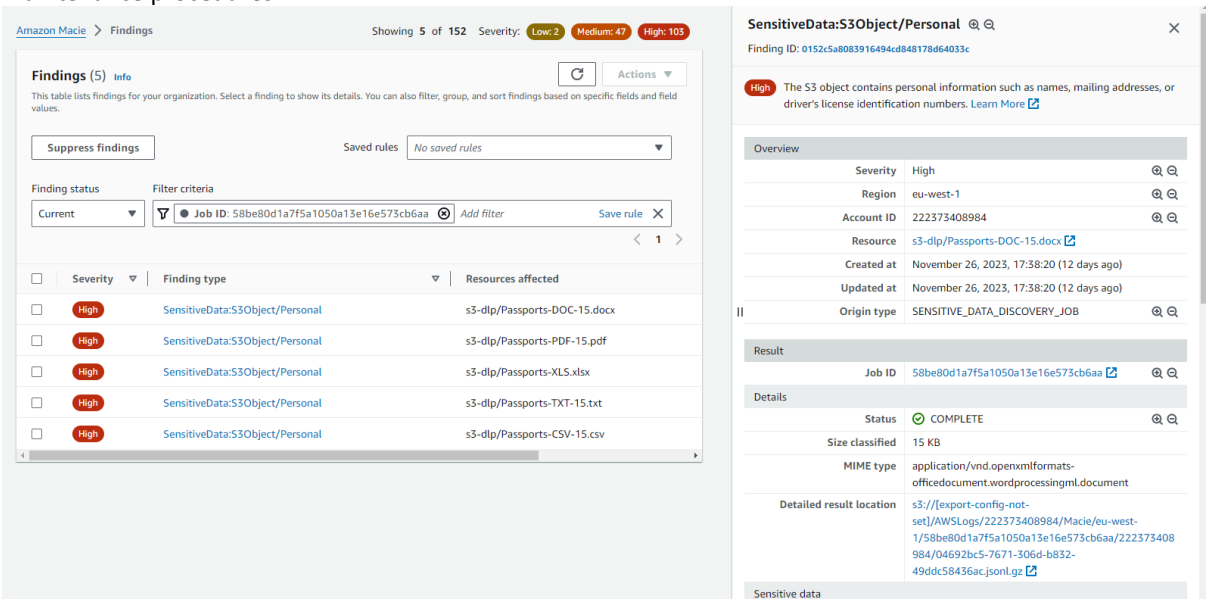
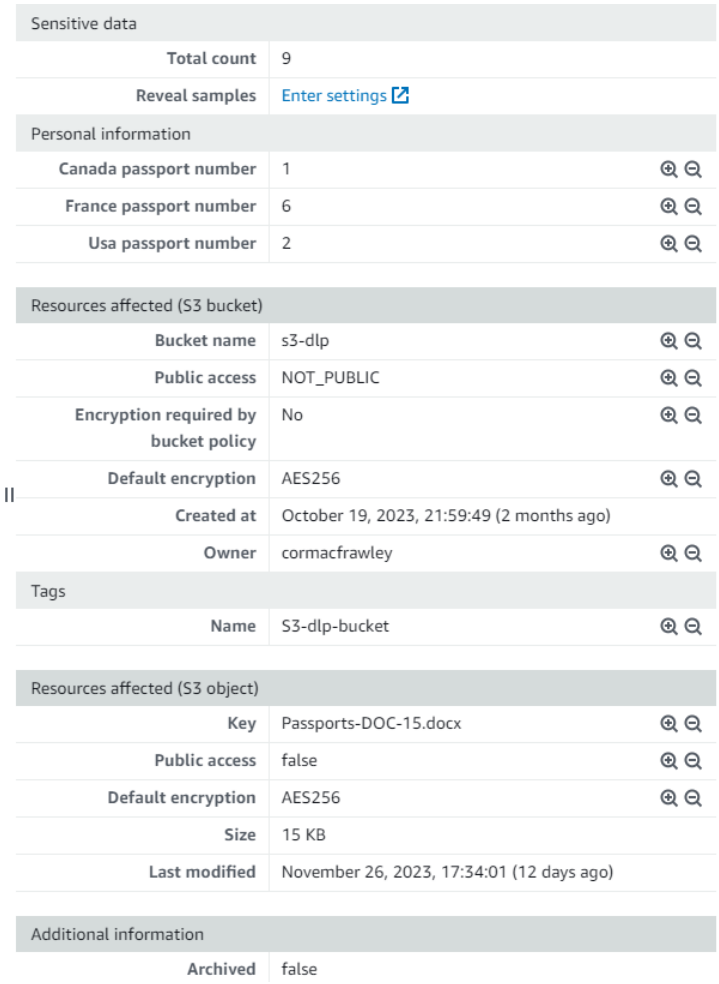



Figure 6 – AWS Macie policy

The results or findings can also be checked in the portal. The documents that triggered for each particular Macie job or policy are listed. For each document the SITs which triggered and the number of occurrences are given. This information can be exported as json or sent to CloudWatch for analysis as part of operation and maintenance procedures.





| Sensitive data | |
|---|---|
| Total count | 9 |
| Reveal samples | Enter settings |
| **Personal information** | |
| Canada passport number | 1 |
| France passport number | 6 |
| Usa passport number | 2 |
| **Resources affected (S3 bucket)** | |
| Bucket name | s3-dlp |
| Public access | NOT_PUBLIC |
| Encryption required by bucket policy | No |
| Default encryption | AES256 |
| Created at | October 19, 2023, 21:59:49 (2 months ago) |
| Owner | cormacfrawley |
| **Tags** | |
| Name | S3-dlp-bucket |
| **Resources affected (S3 object)** | |
| Key | Passports-DOC-15.docx |
| Public access | false |
| Default encryption | AES256 |
| Size | 15 KB |
| Last modified | November 26, 2023, 17:34:01 (12 days ago) |
| **Additional information** | |
| Archived | false |

Figure 7 – AWS Macie policy findings

# 2 GCP

## 2.1 Terraform Code

All Terraform code was written and executed in Visual Studio Code

### 2.1.1 providers.tf

The below code is used to form an authenticated connection to GCP so that the infrastructure and Sensitive Data Protection function can be configured. Also indicates where stored credentials are located. In this instance the credentials are stored as a json file on the local directory.

```
#Provider details
provider "google" {
    #google-beta = google-beta.dev
    # gcloud auth application-default login
    project = "cmacsaff"
    region = "europe-west1"
    credentials = "c:/Cloud_DLP_Thesis/GCP/cmacsaff-76c7db658b5c.json"
}
```

Figure 8 – GCP providers.tf code

### 2.1.2 main.tf

The main.tf file is the main configuration file and the Terraform code executed within it brings up the required GCP cloud environment, storage element (GCP bucket), networking and adds the test files to the bucket.

All storage infrastructure (GCP bucket) is provisioned in the main.tf file along with code to upload files from a directory into that storage. The required infrastructure needed to test is listed below:

- GCP standard storage bucket
- Additional GCP bucket for storing logs
- BigQuery dataset which is a container for generated data
- BigQuery table which is used to stored the data in rows similar to a database
- Object – the object is used to automatically upload files to the storage bucket

```
#GCP storage creation

#Bucket creation
resource "google_storage_bucket" "cmac_dlp" {
    name = "cmac_dlp_bucket"
    location      = "europe-west1"
    storage_class = "STANDARD"

    uniform_bucket_level_access = true

    logging {
    log_bucket = "cmac_dlp_logging_bucket"
  }
}

resource "google_storage_bucket" "cmac_logging" {
    name = "cmac_dlp_logging_bucket"
    location      = "europe-west1"
    storage_class = "STANDARD"

    uniform_bucket_level_access = true


}

#Add files to google bucket
resource "google_storage_bucket_object" "object2" {

    bucket = google_storage_bucket.cmac_dlp.name
    for_each = fileset("c:/Cloud_DLP_Thesis/AWS/uploads/", "*")
    name = each.value
    source = "c:/Cloud_DLP_Thesis/AWS/uploads/${each.key}"

}

#Create a bigquery dataset
resource "google_bigquery_dataset" "source_dataset" {
  dataset_id                  = "cmac_dlp_dataset"
```

```
  description                 = "Input BQ Dataset"
  location                    = "europe-west1"
  default_table_expiration_ms = 3600000


}

#Create a bigquery table
resource "google_bigquery_table" "default" {
  dataset_id = google_bigquery_dataset.source_dataset.dataset_id
  table_id   = "cmac_dlp_dataset_table"
  deletion_protection = false

  time_partitioning {
    type = "DAY"
  }

  labels = {
    env = "default"
    automation = "true"
  }


}
```

Figure 9 – GCP main.tf code

### 2.1.3 dlp.tf

GCP Sensitive Data Protection capabilities are enabled through the dlp.tf code. A basic DLP template is created which aligns to a GCP account and states the SITs which should form part of the policy. Unlike AWS all SITs can be configured through the Terraform code including custom identifiers using regex. Next a trigger for the policy is required which states the scanning interval, the BigQuery dataset and the bucket to be scanned and finally the files types to be included (GCP 2023).

```
#Create a DLP template to search for DLP data
resource "google_data_loss_prevention_inspect_template" "basic" {
    parent          = "projects/cmacsaff"
    description     = "Basic user data identification"
    display_name    = "Basic User Data"

    inspect_config {

        # custom_info_types {
        #     info_type {
        #         name = "CUSTOM_TYPE"
        #     }
        #         likelihood = "UNLIKELY"
        #         regex {
        #             pattern = "^C.+:\\s[0-9]{6}.[0-9]{3}"
        #     }
```

```
    # }

    #   info_types {
    #     name = "EMAIL_ADDRESS"
    # }

    #   info_types {
    #     name = "DATE_OF_BIRTH"
    # }

    #   info_types {
    #     name = "PERSON_NAME"
    # }

    info_types {
        name = "AWS_CREDENTIALS"
    }

    info_types {
        name = "ENCRYPTION_KEY"
    }

    info_types {
        name = "BASIC_AUTH_HEADER"
    }
    info_types {
        name = "HTTP_COOKIE"
    }

      info_types {
        name = "JSON_WEB_TOKEN"

    }

    #   info_types {
    #     name = "UK_NATIONAL_INSURANCE_NUMBER"
    # }


    #   info_types {
    #     name = "VEHICLE_IDENTIFICATION_NUMBER"
    # }


    # info_types {
    #     name = "JSON_WEB_TOKEN"
    # }
```

```
        # info_types {
        #     name = "INDIA_PAN_INDIVIDUAL"
        # }
        # info_types {
        #     name = "AUSTRALIA_TAX_FILE_NUMBER"
        # }
        # info_types {
        #     name = "SPAIN_CIF_NUMBER"
        # }

        # info_types {
        #     name = "BRAZIL_CPF_NUMBER"
        # }

        # info_types {
        #     name = "GERMANY_TAXPAYER_IDENTIFICATION_NUMBER"
        # }

        # info_types {
        #     name = "GERMANY_PASSPORT"
        # }
        # info_types {
        #     name = "FRANCE_PASSPORT"
        # }
        # info_types {
        #     name = "US_PASSPORT"
        # }

        # info_types {
        #     name = "UK_PASSPORT"
        # }

        # info_types {
        #     name = "CANADA_PASSPORT"
        # }



        # info_types {
        #     name = "AUSTRALIA_DRIVERS_LICENSE_NUMBER"
        # }
        # info_types {
        #     name = "SPAIN_DRIVERS_LICENSE_NUMBER"
        # }

        # info_types {
        #     name = "UK_DRIVERS_LICENSE_NUMBER"
        # }
```

```
        # info_types {
        #     name = "JAPAN_DRIVERS_LICENSE_NUMBER"
        # }

        # info_types {
        #     name = "US_DRIVERS_LICENSE_NUMBER"
        # }


        # info_types {
        #     name = "US_SOCIAL_SECURITY_NUMBER"
        # }

        # info_types {
        #     name = "SPAIN_SOCIAL_SECURITY_NUMBER"
        # }
        #  info_types {
        #     name = "CREDIT_CARD_NUMBER"
        # }

        # info_types {
        #     name = "CHINA_RESIDENT_ID_NUMBER"
        # }
        # info_types {
        #     name = "BRAZIL_CPF_NUMBER"
        # }
        # info_types {
        #     name = "ITALY_FISCAL_CODE"
        # }

        # info_types {
        #     name = "INDONESIA_NIK_NUMBER"
        # }

        # info_types {
        #     name = "CROATIA_PERSONAL_ID_NUMBER"
        # }

        # info_types {
        #     name = "INDIA_AADHAAR_INDIVIDUAL"
        # }

        # info_types {
        #     name = "FRANCE_CNI"
        # }

        # info_types {
        #     name = "CHILE_CDI_NUMBER"
        # }
```

```
        min_likelihood = "LIKELY"
        # rule_set {
        #     info_types {
        #         name = "EMAIL_ADDRESS"
        #     }
        #     rules {
        #         exclusion_rule {
        #             regex {
        #                 pattern = ".+@example.com"
        #             }
        #             matching_type = "MATCHING_TYPE_FULL_MATCH"
        #         }
        #     }
        # }

        limits {
            max_findings_per_item    = 1000
            max_findings_per_request = 1000
        }
    }
}

#Create a trigger for the DLP job
resource "google_data_loss_prevention_job_trigger" "basic-bucket-job" {
    parent          = "projects/cmacsaff"
    description     = "Reoccuring search on cmac_dlp_bucket"
    display_name    = "Special-Data Bucket Job"

    triggers {
        schedule {
            recurrence_period_duration = "86400s"
        }
    }

    inspect_job {
        inspect_template_name =
google_data_loss_prevention_inspect_template.basic.id
        actions {
            save_findings {
                output_config {
                    table {
                        project_id = "cmacsaff"
                        dataset_id = "cmac_dlp_dataset"
                    }
                }
            }
        }
        storage_config {
```

```
            cloud_storage_options {
                file_set {
                    url = "gs://cmac_dlp_bucket/"
                }
                file_types = ["TEXT_FILE", "PDF", "CSV", "POWERPOINT", "WORD",
"EXCEL"]
            }
        }
    }
}
```

Figure 10 – GCP dlp.tf code

## 2.2 GCP Sensitive Data Protection policies

Once the GCP policies have been created they can be triggered immediately in the GCP portal and the results examined.

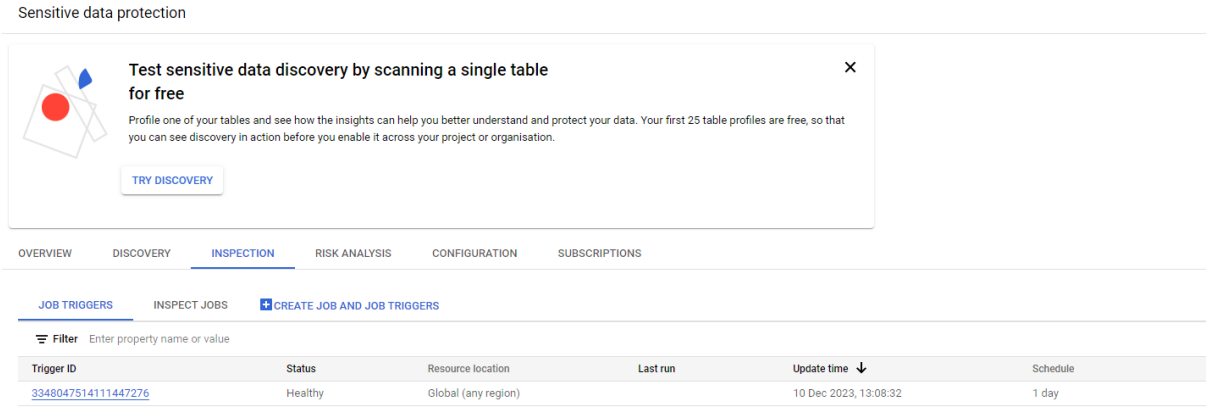



Figure 11 – GCP sensitive data protection

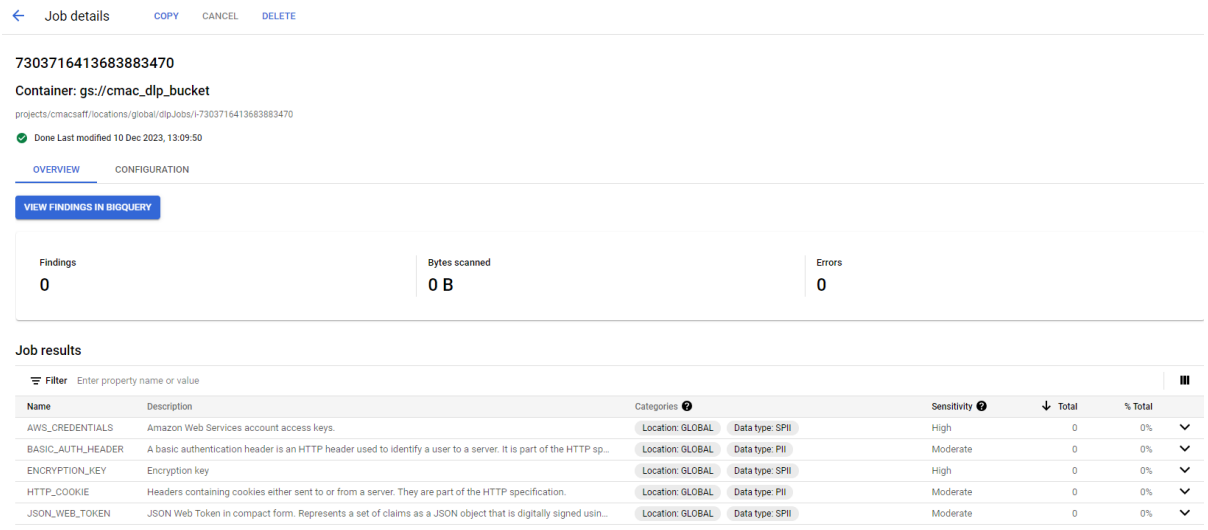



Figure 12 – GCP sensitive data protection findings

# 3 Azure

## 3.1 Terraform Code

All Terraform code was written and executed in Visual Studio Code

### 3.1.1 providers.tf

The below code is used to form an authenticated connection to Azure so that the infrastructure and Sensitive Data Protection function can be configured. For Azure a command line authentication in Visual Studio Code was used utilising the account password so credentials were not included in this file.

```
# Azure Provider source and version being used
terraform {
  required_providers {
    azurerm = {
      source  = "hashicorp/azurerm"
      version = "=3.72.0"
    }
    azapi = {
      source  = "Azure/azapi"
      version = "=1.6.0"
    }
  }
}

# Configure the Microsoft Azure Provider
provider "azurerm" {
  features {}
}
```

Figure 13 – Azure providers.tf code

### 3.1.2 main.tf

The main.tf file is the main configuration file and the Terraform code executed within it brings up the required Azure cloud environment, storage element (Blob Storage), networking and adds the test files to the Blob. Individual SITs could not be configured through Terraform these had to be selected from the Azure portal. The required infrastructure needed to test is listed below:

- A resource group which is a container for all provisioned infrastructure
- A virtual network
- A subnet
- A storage account as a container for the blob storage
- Access rules for the storage account – limited to local machine
- A container for the blob storage
- Blob storage which combined code to upload files to the blob storage once instantiated

```
#Resource Group
resource "azurerm_resource_group" "cmacdlpgrp" {
  name     = "cmacdlpgrp"
  location = "westeurope"
}

resource "azurerm_virtual_network" "cmac-vnet" {
  name                = "cmac-vnet"
  address_space       = ["10.0.0.0/16"]
```

```
  location            = azurerm_resource_group.cmacdlpgrp.location
  resource_group_name = azurerm_resource_group.cmacdlpgrp.name
}

resource "azurerm_subnet" "cmac-subnet" {
  name                 = "cmac-subnet"
  resource_group_name  = azurerm_resource_group.cmacdlpgrp.name
  virtual_network_name = azurerm_virtual_network.cmac-vnet.name
  address_prefixes     = ["10.0.2.0/24"]
  service_endpoints    = ["Microsoft.Storage"]
}

resource "azurerm_storage_account" "cmacdlpacct1974-new" {
  name                     = "cmacdlpacct1974"
  resource_group_name      = azurerm_resource_group.cmacdlpgrp.name
  location                 = "westeurope"
  account_tier             = "Standard"
  account_kind             = "BlobStorage"
  account_replication_type = "LRS"
}

resource "azurerm_storage_account_network_rules" "cmac-nw-rules" {
  storage_account_id = azurerm_storage_account.cmacdlpacct1974-new.id

  default_action             = "Allow"
  ip_rules                   = ["127.0.0.1", "37.xx.xx.xx"]
  virtual_network_subnet_ids = [azurerm_subnet.cmac-subnet.id]
  bypass                     = ["Metrics"]
}

resource "azurerm_storage_container" "cmacdlpcontain" {
  name                  = "cmacdlpcontain"
  storage_account_name  = azurerm_storage_account.cmacdlpacct1974-new.name
  container_access_type = "private"
}


resource "azurerm_storage_blob" "cmacdlpblob-new" {

  for_each = fileset("c:/Cloud_DLP_Thesis/AWS/uploads/", "**/*")
  name = each.key
  storage_account_name = azurerm_storage_account.cmacdlpacct1974-new.name
  storage_container_name = azurerm_storage_container.cmacdlpcontain.name
  type = "Block"
  source = "c:/Cloud_DLP_Thesis/AWS/uploads/${each.key}"

}
```

Figure 14 – main.tf code

### 3.1.3 dlp.tf

DLP.tf enabled the Defender for Cloud modules:

- Defender for Cloud - Storage
- Defender for Cloud – Cloud Security Posture Management

SITs must be defined in the portal and cannot be assigned to policy individually (Azure 2023).

```
data "azurerm_subscription" "current" {}

resource "azurerm_security_center_subscription_pricing" "mdc_storage" {
  tier          = "Standard"
  resource_type = "StorageAccounts"
  subplan       = "DefenderForStorageV2"

     extension {
    name = "SensitiveDataDiscovery"
  }


}

resource "azurerm_security_center_subscription_pricing" "mdc_cspm" {
  tier          = "Standard"
  resource_type = "CloudPosture"

  extension {
    name = "SensitiveDataDiscovery"
  }

}
```

Figure 15 – Azure dlp.tf code

## 3.2 Azure Sensitive Data Protection policies

As mentioned above individual SITs cannot be configured through the Terraform code and are assigned to policy through the portal. All selected SITs are then added to the DLP policy.

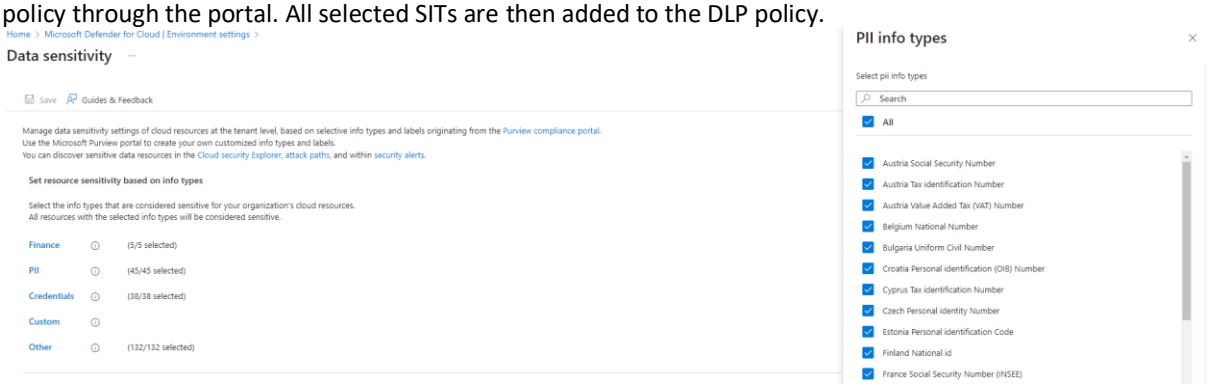



Figure 16 – Azure SIT configuration

Once configured there is no way to check the policies in Azure Defender for Cloud. When the policy triggers an alert is generated which outlines which SITs have been found and in what file types. The exact file is not given.

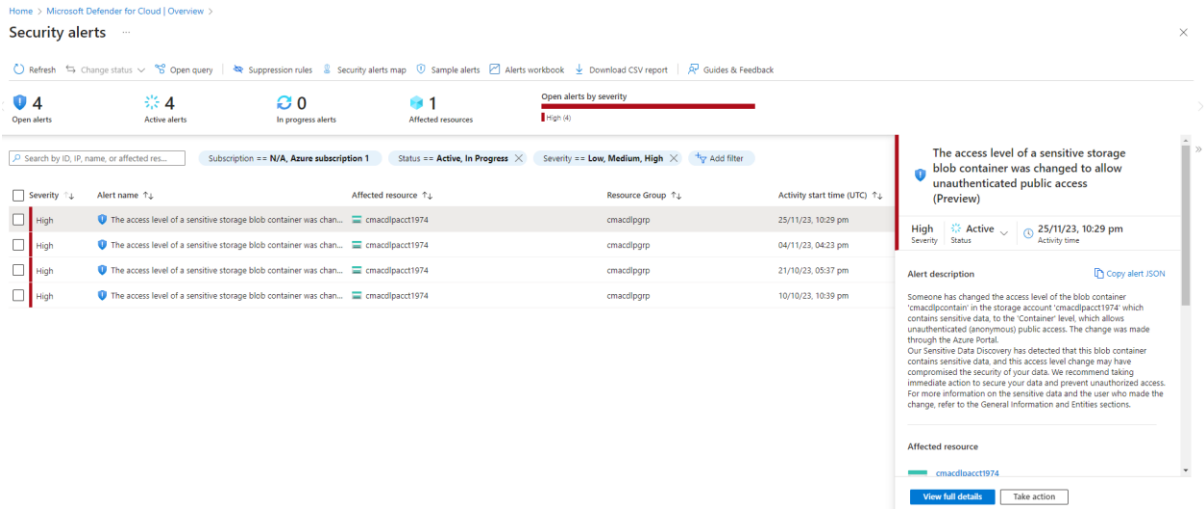



Figure 17 – Azure Alerts

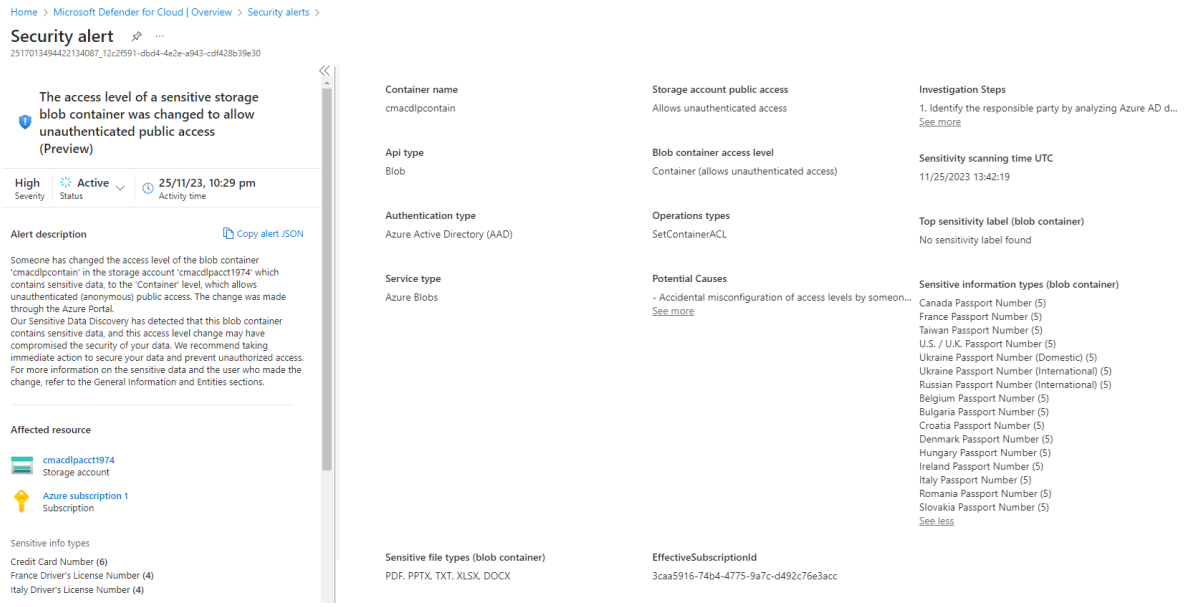



Figure 18 – Azure Sensitive Data Protection findings

# 4 Sensitive Information Type Comparison

The comparison of all available Sensitive Information Types across the three vendors is represented in the table below. The SITs are coloured coded using the below key:

| |
|---|
| Credentials/IT |
| Global |
| India |
| N.AMER |
| LATAM |
| APAC |
| EMEA |

| Managed Data Identifie+A1:D69rs | AWS | GCP | Azure (DFC) |
|---|---|---|---|
| Credentials/IT | Proximity key words required | Key words may be in detector ref | Keywords Incld. |
| Amazon S3 Client Secret Access Key | X | X | Y |
| ASP.NET Machine Key | X | X | Y |
| Azure AD Client Access Token | X | X | Y |
| Azure AD Client Secret | X | X | Y |
| Azure AD User Credentials | X | X | Y |
| Azure App Service Deployment Password | X | X | Y |
| Azure Batch Shared Access Key | X | X | Y |
| Azure Bot Framework Secret Key | X | X | Y |
| Azure Bot Service App Secret | X | X | Y |
| Azure Cognitive Search API Key | X | X | Y |
| Azure Cognitive Service Key | X | X | Y |
| Azure Container Registry Access Key | X | X | Y |
| Azure COSMOS DB Account Access Key | X | X | Y |
| Azure Databricks Personal Access Token | X | X | Y |
| Azure DevOps App Secret | X | X | Y |
| Azure DevOps Personal Access Token | X | X | Y |
| Azure EventGrid Access Key | X | X | Y |
| Azure Function Master / API Key | X | X | Y |
| Azure IoT Shared Access Key | X | X | Y |
| Azure Logic App Shared Access Signature | X | X | Y |
| Azure Machine Learning Web Service API Key | X | X | Y |
| Azure Maps Subscription Key | X | X | Y |
| Azure Redis Cache Connection String | X | X | Y |
| Azure Service Bus connection string | X | X | Y |
| Azure Shared Access Key / Web Hook Token | X | X | Y |
| Azure SignalR Access Key | X | X | Y |
| Azure SQL Connection String | X | X | Y |
| Azure Storage Account Access Key | X | X | Y |
| Azure Storage Account Shared Access Signature | X | X | Y |
| Azure Storage Account Shared Access Signature for High Risk Resources | X | X | Y |
| Azure Subscription Management Certificate | X | X | Y |
| Client Secret / API Key | X | X | Y |
| GitHub Personal Access Token | X | X | Y |
| Google API key | X | X | Y |
| Http Authorization Header | X | X | Y |
| Microsoft Bing Maps Key | X | X | Y |
| Slack Access Token | X | X | Y |
| User Login Credentials | X | Y | X |
| AWS secret access key | Y | Y | X |
| Azure Auth Token | X | Y | X |
| GCP API Key | Y | Y | X |
| GCP Credentials | X | Y | X |
| HTTP Basic Authentication header | Y | Y | X |
| Jason Web Token (JWT) | Y | Y | X |
| XSRF_TOKEN | X | Y | X |
| Password | X | Y | X |
| Weak Password Hash | X | Y | X |
| Auth Token (Oauth/Bearer) | X | Y | X |
| Oauth client secret | X | Y | X |
| Encryption Key (symmetric) | X | Y | X |
| OpenSSH private key | Y | Y | X |
| PGP private key | Y | Y | X |
| Public-Key Cryptography standard (PKCS) private key | Y | Y | X |
| PuTTY private key | Y | Y | X |
| Advertising_ID (online user tracking) | X | Y | X |
| Domain name | X | Y | X |
| ICCID (Sim card ID) | X | Y | X |
| IMEI | X | Y | X |
| IMSI | X | Y | X |
| IP Address (v4 & v6) | X | Y | Y |
| MAC Adress | X | Y | X |
| MAC Address local | X | Y | X |
| Storage Signed policy document | X | Y | X |
| Storage Signed URL | X | Y | X |
| URL | X | Y | X |
| SSL Cert (X.509) | X | Y | Y |
| HTTP Cookie | Y | Y | X |

| Financial Information | | | |
|---|---|---|---|
| IBAN | Y | Y | Y |
| Bank account number - Australia | X | X | Y |
| Bank account number - Canada | Y | Y | Y |
| Bank account number - Japan | X | Y | Y |
| Bank account number - US | Y | X | Y |
| Bank account number - France | Y | X | X |
| Bank account number - Germany | Y | X | X |
| Bank account number - Italy | Y | X | X |
| Bank account number - Israel | X | X | Y |
| Bank account number - New Zealand | X | X | Y |
| Bank account number - Spain | Y | X | X |
| Bank account number - UK | Y | X | X |
| Credit card expiry date | Y | X | X |
| Credit card magnetic stripe data | Y | Y | X |
| Credit card track number | X | Y | X |
| Credit card number | Y | Y | Y |
| Credit card verification code | Y | Y | X |
| EU Debit Card | X | X | Y |
| SWIFT Code (BIC) | X | Y | Y |
| Portugal NIB (Bank ID no) | X | Y | X |
| US CUSIP ID (Bank ID no) | X | Y | X |
| US Bank Routing MICR (ABA/RTN) | X | Y | Y |
| VAT Number - Austria | X | X | Y |
| VAT Number - Belgium | X | Y | Y |
| VAT Number - France | X | Y | X |
| VAT Number - Germany | X | Y | Y |
| VAT Number - Hungary | X | Y | Y |
| VAT Number - Indonesia | X | Y | X |
| VAT Number - Italy | X | Y | Y |
| VAT Number - Netherlands | X | Y | Y |
| Tax identification or reference number - Australia | Y | Y | Y |
| Tax identification or reference number - Austria | x | x | Y |
| Tax identification or reference number - Brazil (CPF) | Y | Y | Y |
| Tax identification or reference number - Brazil (CNPJ) | Y | X | Y |
| Tax identification or reference number - Cyprus | X | Y | Y |
| Tax identification or reference number - France | Y | Y | Y |
| Tax identification or reference number - Germany | Y | Y | Y |
| Tax identification or reference number - Greece | X | X | Y |
| Tax identification or reference number - Hungary | X | X | Y |
| Tax identification or reference number - Italy | Y | X | Y |
| Tax identification or reference number - India (PAN) | Y | Y | Y |
| Tax identification or reference number - Malta | X | X | Y |
| Tax identification or reference number - Netherlands | X | X | Y |
| Tax identification or reference number - Norway | X | Y | X |
| Tax identification or reference number - New Zealand (IRD) | Y | Y | Y |
| Tax identification or reference number - Poland | X | X | Y |
| Tax identification or reference number - Portugal | X | X | Y |
| Tax identification or reference number - Slovenia | X | X | Y |
| Tax identification or reference number - Spain (CIF) | Y | Y | X |
| Tax identification or reference number - Spain (NIF) | X | Y | Y |
| Tax identification or reference number - Sweden | X | X | Y |
| Tax identification or reference number - UK | X | Y | Y |
| Tax identification or reference number - US | X | Y | Y |
| US Adoption Taxpayer ID | X | Y | X |
| Germany Schufa ID | X | Y | X |
| US Employer ID Number (EIN) | X | Y | X |
| US Individual Taxpayer Identification Number (ITIN) | X | Y | X |

| Personal Information - PII | | | |
|---|---|---|---|
| DOB | Y | Y | X |
| Full Name/Person Name | Y | Y | X |
| Female name | X | Y | X |
| Male Name | X | Y | X |
| Last Name | X | Y | X |
| Age | X | Y | X |
| Country Demographic | X | Y | X |
| Email | X | Y | X |
| Ethnic Group | X | Y | X |
| Gender | X | Y | X |
| Generic ID | X | Y | X |
| Marital Status | X | Y | X |
| Drivers License ID number - Australia | Y | Y | Y |
| Drivers License ID number - Austria | Y | X | Y |
| Drivers License ID number - Belgium | Y | X | Y |
| Drivers License ID number - Bulgaria | Y | X | Y |
| Drivers License ID number - Canada | Y | Y | Y |
| Drivers License ID number - Croatia | Y | X | Y |
| Drivers License ID number - Cyprus | Y | X | Y |
| Drivers License ID number - Czech Rep | Y | X | Y |
| Drivers License ID number - Denmark | Y | X | Y |
| Drivers License ID number - Estonia | Y | X | Y |
| Drivers License ID number - Finland | Y | X | Y |
| Drivers License ID number - France | Y | X | Y |
| Drivers License ID number - Germany | Y | Y | Y |
| Drivers License ID number - Greece | Y | X | Y |
| Drivers License ID number - Hungary | Y | X | Y |
| Drivers License ID number - Ireland | Y | Y | Y |
| Drivers License ID number - India | Y | X | X |
| Drivers License ID number - Italy | Y | X | Y |
| Drivers License ID number - Japan | X | Y | Y |
| Drivers License ID number - Latvia | Y | X | Y |
| Drivers License ID number - Lithuania | Y | X | Y |
| Drivers License ID number - Luxembourg | Y | X | Y |
| Drivers License ID number - Malta | Y | X | Y |
| Drivers License ID number - Netherlands | Y | Y | Y |
| Drivers License ID number - New Zealand | X | X | Y |
| Drivers License ID number - Poland | Y | X | Y |
| Drivers License ID number - Portugal | Y | X | Y |
| Drivers License ID number - Romania | Y | X | Y |
| Drivers License ID number - Slovakia | Y | X | Y |
| Drivers License ID number - Slovenia | Y | X | Y |
| Drivers License ID number - Spain | Y | Y | Y |
| Drivers License ID number - Sweden | Y | X | Y |
| Drivers License ID number - UK | Y | Y | Y |
| Drivers License ID number - US | Y | Y | Y |

| | | | |
|---|---|---|---|
| UK electoral roll number | Y | X | Y |
| GPS co-ordinates | Y | X | X |
| Mailing Adress - Australia (certain key words needed in address) | Y | X | X |
| Mailing Address - Brazil | Y | | X |
| Mailing Adress - Canada (certain key words needed in address) | Y | X | X |
| Mailing Adress - France (certain key words needed in address) | Y | X | X |
| Mailing Adress - Germany (certain key words needed in address) | Y | X | X |
| Mailing Adress - Italy (certain key words needed in address) | Y | X | X |
| Mailing Adress - Spain (certain key words needed in address) | Y | X | X |
| Mailing Adress - UK (certain key words needed in address) | Y | X | X |
| Mailing Adress - US (certain key words needed in address) | Y | X | X |
| Location Physical address | X | Y | X |
| Street Address | X | Y | X |
| Ireland - eircode | X | Y | X |
| National Identification Number - Argentina (DNI) | X | X | Y |
| National Identification Number - Austria | X | X | Y |
| National Identification Number - Belgium | X | X | Y |
| National Identification Number - Brazil | Y | X | Y |
| National Identification Number - Bulgaria | X | X | Y |
| National Identification Number - Chile | X | Y | Y |
| National Identification Number - China | X | Y | Y |
| National Identification Number - Colombia | X | Y | X |
| National Identification Number - Croatia (OIB & ID card) | X | Y | Y |
| National Identification Number - Cyprus | X | Y | Y |
| National Identification Number - Czech | X | Y | Y |
| National Identification Number - Denmark | X | Y | Y |
| National Identification Number - Estonia | X | X | Y |
| National Identification Number - France CNI | Y | Y | Y |
| National Identification Number - Finland | X | X | Y |
| National Identification Number - Germany | Y | Y | Y |
| National Identification Number - Greece | X | X | Y |
| National Identification Number - Hong Kong | X | Y | Y |
| National Identification Number - Hungary | X | X | Y |
| National Identification Number - Italy | Y | Y | X |
| National Identification Number - India (Aadhaar) | Y | Y | Y |
| National Identification Number - Indonesia | X | Y | Y |
| National Identification Number - Israel | X | Y | Y |
| National Identification Number - Japan (residence card no) | X | Y | Y |
| National Identification Number - Japan (residence registration number) | X | Y | Y |
| National Identification Number - Japan (My Number) | X | Y | Y |
| National Identification Number - Korea | X | X | Y |
| National Identification Number - Latvia | X | X | Y |
| National Identification Number - Lithuania | X | X | Y |
| National Identification Number - Luxemburg | X | X | Y |
| National Identification Number - Malta | X | X | Y |
| National Identification Number - Malaysia | X | X | Y |
| National Identification Number - Mexico | X | Y | X |
| National Identification Number - Netherlands | X | Y | Y |
| National Identification Number - Norway | X | X | Y |
| National Identification Number - Paraguay | X | Y | X |
| National Identification Number - Peru | X | Y | X |
| National Identification Number - Poland (PSEL & National ID card) | X | Y | Y |
| National Identification Number - Poland (REGON) | X | X | Y |
| National Identification Number - Portugal | X | Y | Y |
| National Identification Number - Philippines | X | X | Y |
| National Identification Number - Romania | X | X | Y |
| National Identification Number - Saudi Arabia | X | X | Y |
| National Identification Number - Slovakia | X | X | Y |
| National Identification Number - Slovenia | X | X | Y |
| National Identification Number - Singapore | X | Y | Y |
| National Identification Number - Spain | Y | Y | Y |
| National Identification Number - South Africa | X | Y | Y |
| National Identification Number - Sweden | X | Y | Y |
| National Identification Number - Taiwan ARC/TARC | X | X | Y |
| National Identification Number - Taiwan National ID | X | X | Y |
| National Identification Number - Thailand | X | Y | X |
| National Identification Number - Turkey | X | Y | X |
| National Identification Number - Uruguay | X | Y | X |
| National Identification Number - Venezuela | X | Y | X |

| | | | |
|---|---|---|---|
| Passport Number - Australia | X | Y | Y |
| Passport Number - Austria | X | X | Y |
| Passport Number - Belgium | X | X | Y |
| Passport Number - Bulgaria | X | X | Y |
| Passport Number - Canada | Y | Y | Y |
| Passport Number - China | X | Y | X |
| Passport Number - Croatia | X | X | Y |
| Passport Number - Cyprus | X | X | Y |
| Passport Number - Czech | X | X | Y |
| Passport Number - Denmark | X | X | Y |
| Passport Number - Estonia | X | X | Y |
| Passport Number - Finland | X | X | Y |
| Passport Number - France | Y | Y | Y |
| Passport Number - Germany | Y | Y | Y |
| Passport Number - Greece | X | X | Y |
| Passport Number - Hungary | X | X | Y |
| Passport Number - Italy | Y | X | Y |
| Passport Number - Ireland | X | Y | Y |
| Passport Number - Japan | X | Y | Y |
| Passport Number - Korea | X | Y | X |
| Passport Number - Latvia | X | X | Y |
| Passport Number - Lithuania | X | X | Y |
| Passport Number - Luxemburg | X | X | Y |
| Passport Number - Malta | X | X | Y |
| Passport Number - Mexico | X | Y | X |
| Passport Number - Netherlands | X | Y | Y |
| Passport Number - Philippines | X | X | Y |
| Passport Number - Poland | X | Y | Y |
| Passport Number - Portugal | X | X | Y |
| Passport Number - Romania | X | X | Y |
| Passport Number - Russia (domestic) | X | X | Y |
| Passport Number - Russia (international) | X | X | Y |

| | | | |
|---|---|---|---|
| Passport Number - Slovakia | X | X | Y |
| Passport Number - Slovenia | X | X | Y |
| Passport Number - Singapore | X | Y | X |
| Passport Number - Spain | Y | Y | Y |
| Passport Number - Sweden | X | Y | Y |
| Passport Number - Taiwan | X | Y | Y |
| Passport Number - UK | Y | Y | Y |
| Passport Number - Ukraine Domestic | X | X | Y |
| Passport Number - Ukraine Internaltional | X | X | Y |
| Passport Number - US | Y | Y | Y |
| Permanent Residence Number - Canada | Y | X | X |
| Phone Number - Brazil | Y | X | X |
| Phone Number - Canada | Y | X | X |
| Phone Number - France | Y | X | X |
| Phone Number - Germany | Y | X | X |
| Phone Number - Italy | Y | X | X |
| Phone Number - Spain | Y | X | X |
| Phone Number - UK | Y | X | X |
| Phone Number - US | Y | X | X |
| Phone Number | Y | Y | X |
| Social Security Number - Austria | x | x | Y |
| Social Security Number - France | x | x | Y |
| Social Security Number - Germany | x | x | X |
| Social Security Number - Greece | x | x | Y |
| Social Security Number - Hungary | x | x | Y |
| Social Security Number - Japan | x | x | Y |
| Social Security Number - Switzerland | x | x | Y |
| Social Security Number - US | Y | Y | Y |
| Social Security Number - Spain | Y | Y | Y |
| Social Security Number - Korea | X | Y | X |
| Social Security Number - Portugal | X | Y | X |
| Social Insurnace Number - Canada | Y | Y | Y |
| Social Insurnace Number - Ireland (PPSN) | X | Y | Y |
| Social Welfare No New Zealand | X | x | Y |
| National Insurnace Number (UK) | Y | Y | Y |
| Vechile Identificatin number (VIN) | Y | Y | X |
| **Other** | | | |
| Date | X | Y | X |
| Organisation Name | X | Y | X |
| Time | X | Y | X |
| India GST Individual (Business ID) | X | Y | X |
| US State | X | Y | X |
| US Toll free number | X | Y | X |
| Australia Company Number | X | X | Y |
| Australia Business Number | X | X | Y |
| Japan My Number Corporate | X | X | Y |

# 5 Testing Results

All results of the testing for AWS & GCP are listed in the below table – Y successful, N failed and P partial. Those tests that were tagged as partial were as a result of a successful trigger of the policy but where the correct number of instances of the SIT was not given. Azure testing is not included as the functionality did not allow for a like for like comparison.

| SIT | AWS | | | | | GCP | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | DOC | PDF | XLSX | CSV | TXT | DOC | PDF | XLSX | CSV | TXT | PPT |
| Credit Card Number (5) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| IBAN GB (2) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| IBAN ES (2) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| IBAN DE (2) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| IBAN IT (1) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| SSN US (10) | P | P | Y | Y | P | P | P | Y | Y | P | P |
| SSN ES (2) | N | N | Y | Y | N | N | N | N | N | N | N |
| PII Name | Y | Y | Y | Y | Y | Y | Y | Y | P | P | Y |
| PII DOB | N | N | Y | Y | N | N | N | P | Y | N | N |
| PII Email | | | | | | Y | Y | Y | Y | P | Y |
| Driving License US | P | P | P | P | P | P | P | P | P | P | P |
| Driving License UK* | N | N | Y | Y | N | N | N | Y | N | N | N |
| DrivingLicense Australia | N | N | Y | Y | N | N | N | N | N | N | N |
| Driving License Japan | | | | | | N | N | N | N | N | N |
| Driving License Italy | Y | N | Y | Y | N | | | | | | |
| Driving License France | N | N | Y | Y | N | | | | | | |
| Driving License Spain* | N | N | N | N | N | N | N | N | N | N | N |
| NIN India (AADHAAR) | N | N | Y | Y | N | Y | Y | P | P | Y | Y |
| NIN France (CNI) | N | N | N | N | N | N | N | N | N | N | N |
| NIN Brazil (CPF) | N | N | Y | N | N | Y | Y | Y | N | Y | Y |
| NIN Brazil (RG) | N | N | Y | Y | N | | | | | | |
| NIN Italy (Fiscal Code) | N | N | N | N | N | Y | Y | Y | Y | Y | Y |
| NIN Indonesia (NIK) | | | | | | N | N | N | N | N | N |
| NIN Croatia (OIB) | | | | | | Y | Y | N | N | Y | Y |
| NIN China (Resident ID No) | | | | | | N | N | N | N | N | N |
| NIN Chile (CDI) | | | | | | Y | Y | Y | Y | Y | Y |
| Passport UK | N | Y | Y | Y | Y | Y | Y | N | N | Y | Y |
| Passport US | N | N | Y | Y | N | Y | Y | N | N | Y | Y |
| Passport France | Y | N | Y | Y | N | P | P | N | N | P | P |
| Passport Germany | N | N | N | N | N | P | P | N | N | P | P |
| Passport Canada | N | N | Y | Y | N | Y | Y | N | N | Y | Y |
| Tax No Australia | N | N | Y | Y | N | Y | Y | Y | Y | Y | Y |
| Tax No Brazil (CPF) | N | N | Y | Y | N | N | N | Y | Y | N | N |
| Tax No Germany | N | N | Y | Y | N | N | N | N | N | N | N |
| Tax No India | | | | | | Y | Y | Y | Y | Y | Y |
| UK National Insurance Number | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y |
| AWS secret key | N | N | N | N | N | N | N | N | N | N | N |
| Putty private key | N | N | N | N | N | | | | | | |
| OpenSSH private Key | Y | N | N | N | Y | | | | | | |
| PGP private key | Y | N | N | N | Y | | | | | | |
| Encryption Key | | | | | | P | N | N | N | N | P |
| HTTP Basic Authentication header | Y | Y | Y | Y | Y | N | N | N | N | N | N |
| Jason Web Token (JWT) | N | N | N | N | N | Y | N | Y | Y | Y | Y |
| Vechile Identification Number | N | N | N | N | N | Y | Y | Y | Y | Y | Y |
| Custom 1 | Y | | | | | N | | | | | |

# References

AWS (2023) *What is Amazon Macie?* Available at: https://docs.aws.amazon.com/macie/latest/user/what-is-macie.html [Accessed 18 November 2023]

GCP (2023) *Sensitive Data Protection overview* Available at: https://cloud.google.com/dlp/docs/sensitive-data-protection-overview [Accessed 18 November 2023]

Microsoft (2023) *What is Microsoft Defender for Cloud?* Available at: https://learn.microsoft.com/en-us/azure/defender-for-cloud/defender-for-cloud-introduction [Accessed 18 November 2023]