DMARC XML to CSV

Ceci est une documentation du script Powershell **DMARC-XML_to_CSV.** Ce script sert à décompresser et extraire l'ensemble des rapports DKIM non-valides de rapports XML compressés. Les rapports filtrés sont ensuite enregistrés en CSV, et regroupés par domaines concernés.

Sommaire

1. Installation

Ce script n'utilise uniquement des commandes Powershell. Il n'y a pas d'installation externe à effectuer.

DMARC-XML_to_CSV.ps1

2. Utilisation

1. Lancer le script sur l'Invite de Commandes ou sur Powershell avec une des deux commandes:
    - `powershell -F <chemin complet du script>`
    - `./<script>`
2. Donner le chemin entier du dossier contenant les fichiers compressés.

📣 Le dossier doit contenir les fichiers XML compressés, et sera le seul dossier modifié.

3. Processus

ℹ️ Etapes du programme:

1. Décompression des fichiers **Zip** et **Gzip** dans le chemin spécifié.
2. Récupère les fichiers **XML** dans le dossier et les traite **un par uns.**
   Processus suivi par fichier:
   a. Extraction des données principales du fichier en matrice de chaînes de caractères.
   b. Vérifie le nombre d'entrées de la nouvelle matrice.
   c. Effectue la transpose de la matrice.
   d. Trie les entrées et les ajoute à un **fichier CSV**.

   → En cas d'erreur, déplace les fichiers dans un dossier nommé "Exceptions", et passe au fichier suivant.

4. Fonctions

- `Get-DMARC-data` : Permet d'extraire les informations d'un fichier DMAC .xml, et les garde dans une matrice.
- `Get-DMARC-wNS-data` : Copie de la fonction `Get-DMARC-data`, mais modifié pour supporter les fichiers avec Namespaces (environ 8 cas sur 166, 0.05%)
- `Check-data-processing` : Vérifie que les données ont été extraites correctement du fichier .xml.
- `Reorganize_DMARC_data` : Réarrange la matrice d'informations: on passe de un <record> par colonne à un <record> par ligne (xy → yx) pour préparer à remplir le fichier CSV.
- `Decompressing_Files` : Décompresse les fichiers Gzip, Zip, et renomme les fichiers sans extension pour avoir une liste de fichiers .xml uniquement.

- `Fill-DMARC-csv` : Filtre les <records> de la matrice donnée, et garde uniquement celles avec une mention de 'fail'. Crée ensuite un fichier CSV(FR) avec ces <records> spécifiques.
- `Main` : Fonction principale du document pour appeler les autres fonctions.

5. Description des données

Un document XML DMARC est un rapport généré contenant les analyses de sécurité des mails envoyés depuis un domaine. Il contient de nombreuses informations, comme les informations d'envoi, les authentifications de domaine *(qui permettent de déterminer l'authenticité d'un mail),* ainsi que si le mail a été reçu, rejeté, ou encore considéré comme du spam.

- **Informations du rapport :**
  - **Origine du rapport :** Correspond à la source du fichier XML; le **fournisseur de services de messagerie** ayant envoyé le fichier XML d'origine.
  - **Date de début et de fin :** (UTC, nous sommes en UTC+1). Correspondent à la **période analysée** dans le fichier XML d'origine. *Les mails peuvent avoir été envoyés n'importe quand pendant cette période.*

- **Informations d'envoi :**
  - **IP Source** : Adresse IP de l'expéditeur.
  - **Destination :** Domaine de destination *(TO: exemple@<destination>).*
  - **Domaine Source:** Domaine d'origine *(FROM: exemple@<domaine_source>).*
  - **Return-Path :** Adresse définie vers laquelle sera redirigée emails d'erreurs et autres messages automatiques *(messages de rejet, réponses automatiques d'absences, etc. Les réponses directes ne sont pas redirigées).*

- **Politique DMARC :** Vérifie que les authentifications via SPF et/ou DKIM sont valides *(ce qui prouve l'authenticité des mails)* et indique le traitement du mail choisi en conséquence.
    - **SPF Match :** Vérifie que le domaine SPF fourni correspond bien au domaine DMARC défini dans la politique du document *(le domaine en question se trouve dans le nom du fichier CSV)*.
    - **DKIM Match :** Vérifie que un des domaines DKIM fournis correspondent bien au domaine DMARC défini dans la politique du document *(le domaine en question se trouve dans le nom du fichier CSV)*.
    - **DMARC Resultat:** Indique l'action prise vis-à-vis du mail.
        - **none** : aucune action prise; le mail est passé.
        - **quarantine :** le mail est considéré comme suspect, donc reçu mais dans la catégorie spam.
        - **reject** : le mail est rejeté.

- **Politique DKIM** : Un protocole d'authentification qui agit en tant que signature numérique *(système de clés privées/publiques asymétriques)*, permet d'indiquer que le mail vient bien de son domaine d'origine. Permet aussi de garantir qu'un email n'a pas été altéré avant réception.

❓ ▼ ***Mais encore?***

Une entreprise possédant son propre domaine pour ses mails possède deux clés *(en général appelées publique et privée)*, qui lui permettent de chiffrer les informations et de garantir intégrité et confidentialité. Les deux clés peuvent déchiffrer les informations chiffrées par l'autre, et ce dans n'importe quel sens, ce qui permet d'avoir un échange sûr.
Le DKIM est un protocole qui vérifie bien que le mail est seulement déchiffrable par la clé publique du domaine d'origine, ce qui implique que le mail ait été chiffré par l'autre clé, la clé privée, qui est strictement confidentielle au domaine. Si la vérification est correcte, cela permet de confirmer que le mail est bel et bien envoyé par son domaine spécifié.

- **DKIM Resultat:** Décrit le résultat d'analyse des différents DKIM d'un mail.
  - **pass** : signature valide, email non-altéré.
  - **fail** : signature non-valide, peut venir d'une source non validée ou être un email altéré.
  - **neutral** : peut arriver s'il n'y a pas de DKIM, ou s'il y a eu une erreur.
  - **temperror** : erreur temporaire *(erreur de résolution DNS, etc),* qui peut passer en "pass" si une revérification est effectuée.
  - **permerror** : erreur permanente, désigne une erreur de configuration DKIM, un enregistrement DNS incorrect ou encore un format de signature invalide.
- **DKIM Domaine** : domaine de référence pour la signature DKIM. Contient également la clé publique de la signature.
- **DKIM Selector** : permet de localiser les clés privées utilisées lors de la signature DKIM.

- **Politique SPF**: Un protocole d'authentification qui permet d'indiquer quelles adresses IP de serveurs de messageries sont autorisées à envoyer un mail

sous le nom d'un domaine spécifique.

❓ ▼ ***Mais encore?***

Une entreprise possédant son propre domaine pour ses mails va préciser les plages IP pouvant envoyer des mails sous son domaine/nom.
La politique SPF vérifie que l'adresse IP de l'expéditeur fasse bien partie des plages IP autorisées par le domaine spécifié.

- **SPF Resultat:** Décrit le résultat d'analyse du SPF d'un mail.
  - **pass** : adresse IP valide, et autorisée par le domaine.
  - **fail** : adresse IP non-valide, vient d'une source non validée.
  - **softfail** : adresse IP non explicitement autorisée, mais considéré comme 'tolérée'.
  - **temperror** : erreur temporaire *(erreur de résolution DNS, etc)*, qui peut passer en "pass" si une revérification est effectuée.
  - **permerror** : erreur permanente, désigne une erreur de configuration SPF ou un enregistrement DNS incorrect.
- **SPF Domaine** : domaine de référence au protocole à consulter pour vérifier la liste d'adresses IP acceptées.