



Le renseignement d'origine sources ouvertes



Manuel Medina
Analyste de renseignement

Les initiatives d'anti-corruption, de la transparence et de la liberté de l'information sur ces dernières décennies ont significativement augmentées la valeur des renseignements d'origine sources ouvertes pour les secteurs privés et publics.

Dans ce guide rapide, l'analyste de renseignement Manuel Medina explique le renseignement d'origine sources ouvertes et explore des questions difficiles inhérentes.

Comment le renseignement d'origine sources ouvertes rentre dans le cycle du renseignement

Le renseignement d'origine sources ouvertes, connu sous l'acronyme ROSO dans le secteur du renseignement, est la collecte, le traitement et l'analyse systématique des informations des sources ouvertes. Cela veut dire l'informations ouvertes au public, tels que les articles de presse, les

rapports de recherche ou les registre des sociétés dans un registre public des entreprises.

Comme je l'ai expliqué dans mon précédent guide rapide sur le renseignement, il est utile de voir le renseignement comme un processus au lieu d'un objet fixe. Les données qui sortent de ce processus peuvent être des produits stratégiques ou opérationnels. Dans les deux cas, il est conçu afin d'aider des décideurs à comprendre un sujet spécifique ou une menace, afin de décider comment agir pour atteindre des objectifs spécifiques. Si le renseignement est un processus, l'informations des sources ouvertes sont une des contributions possibles à ce processus.

Seule, l'informations des sources ouvertes peuvent s'avérer un outil puissant afin de soutenir la prise de décision fondée sur des preuves. Selon le contexte et les besoins, les analystes peuvent aussi combiner des informations des sources ouvertes à des informations confidentielles dans le cycle du renseignement.

Par exemple, si un informateur ou un dénonciateur divulgue qu'une personne en particulier a des intérêts d'affaires non-divulgués ou suspects. L'informations des sources ouvertes d'un registre d'entreprise, des sites web et des articles de presse peuvent éventuellement le corroborer.

Le ROSO, peut-il appuyer les services de détection et de répression ?

Oui, c'est même essentiel. Dans de nombreux pays, les rapports des médias publics, surtout ceux des journalistes d'investigation, peuvent déclencher des enquêtes en matière d'infractions financières. Dans certains cas, les personnes se trahissent avec des informations rendues disponibles au public sur les réseaux sociaux. Cela s'est passé ainsi avec l'affaire de la famille Ambuila en Colombie. Leur fille a posté des photos en ligne en possession de biens luxueux et une Lamborghini.

Tel que décrit par mon collègue Tom Walugembe, la première affaire de blanchiment de capitaux en Uganda a débuté par une vidéo WhatsApp largement diffusée qui montrait les criminels prenant la pose à côté des liasses de dollars dans leur appartement. Ce n'est pas très intelligent.

D'où vient le ROSO ?

Nous associons le terme « sources ouvertes » aux informations numériques, mais il peut aussi être physique. Jadis, les analystes et les enquêteurs étaient obligés de se rendre aux bibliothèques et aux bureaux gouvernementaux afin d'obtenir des documents et des articles publics. Vous avez déjà vu des films où la police fouille dans des tas de vieux journaux dans un local d'archives poussiéreux ?

De nos jours, beaucoup de ces informations sont disponibles sur internet et les analystes passent bien plus de temps derrière leurs bureaux. Mais il

existe dans le monde beaucoup d'informations des sources ouvertes qui sont toujours sur papier, surtout dans les pays ayant des systèmes informatiques moins développés.

Il faut noter que « disponible » n'est pas synonyme de « gratuit ». Beaucoup de journaux et d'articles de presse se trouvent derrière des restrictions d'accès payant, par exemple. Payer pour des informations disponibles à tous est une question ambiguë et franchement un casse-tête pour des équipes de renseignement ayant peu de ressources.

Tant que le coût n'est pas prohibitif et qu'il n'existe pas d'autres restrictions à l'accès, généralement nous classifions ces données en sources ouvertes.

La valeur du ROSO pour les secteurs privés et publics

Des renseignements d'origine sources ouvertes sont précieux aux services de détection et de répression. Ils génèrent des pistes, corroborent des informations reçues d'autres sources et étayent des preuves qui serviront devant un tribunal.

Il est également utile pour des entreprises et des institutions financières dans les évaluations des risques et les mesures de vigilance en matière de clients, d'employés et de tiers. Des sources ouvertes peuvent servir à identifier des réseaux frauduleux ou de contrefaçon, ou à enquêter sur les propos d'un dénonciateur. Généralement, le renseignement d'affaire stratégique est aussi basé sur les données de marché d'origine sources ouvertes.

L'information, égale-t-elle le pouvoir ? Plus maintenant

Depuis quelques décennies, l'expansion rapide d'Internet et des outils de recherche a rendu disponible au public une quantité importante d'information. En quelques clics, nous pouvons trouver des archives de journaux datant de plusieurs années, des enregistrements de jugements des tribunaux, ou des bases de données téléchargeables de statistiques et de mesures.

Paradoxalement, l'accès rapide et facile aux quantités importantes de données rend désuet le vieil adage « l'information, c'est le pouvoir ».

Premièrement, il y a le problème de fausses nouvelles – par exemple, des articles délibérément trompeurs ou des postes sur des réseaux sociaux sans fondement factuel. Les fausses nouvelles qui deviennent virales ajoutent de l'huile sur le (faux) feu. Les analystes doivent suivre des procédures extensives afin de vérifier la fiabilité des informations des sources ouvertes à la collecte et ensuite lors des étapes d'intégration et de traitement.

Deuxièmement, trop d'information peut facilement nous « infotoxiquer » et brouiller les processus décisionnels. Rassembler des données pertinentes et en extraire des tendances utiles est le pouvoir, pas seulement

l'information.

Comment lutter contre « l'infotoxication »

La lutte contre « l'infotoxication » est un problème sérieux. Si vous avez passé des heures à vous perdre dans les dédales d'Internet lors des recherches sur un sujet, vous n'êtes pas le seul. Des outils aidant à accélérer le processus de recherche, extraire des informations pertinentes et relier les points sont d'une utilité précieuse.

Notre outil, le [Basel Open Intelligence](#) est conçu exactement pour cela. Il exécute des recherches automatisées en un clic sur le nom d'un individu ou d'une organisation en combinaison avec plus de 200 mots-clés en matière d'infractions financières, d'intervention judiciaire, ou d'autres infractions criminelles et des listes personnalisées de mots-clés. Il fouille aussi le deep web, à la recherche des références sur des listes de sanctions et de personnes politiquement exposées.

Vous pouvez rechercher de multiples pseudonymes et variations de nom en même temps. Le nombre de variations et de translittérations des noms est étonnant et frustrant, surtout dans les langues aux alphabets différents ou des structures de noms complexes. Certaines personnes utilisent des pseudonymes pour dissimuler leur véritable identité.

L'algorithme utilise le traitement automatique du langage naturel afin d'en extraire des entités et dévoiler des relations entre des personnes, des entreprises, des lieux et des postes professionnels. Une liste de documents trouvés dans la recherche est générée, ainsi que le texte principal extrait du site web en dehors du contenu sans rapport, tels que les publicités, les menus et les notifications de cookies. Les mots-clés sont surlignés pour une lecture rapide. L'utilisateur peut rapidement filtrer et trier les documents selon leur pertinence.

Ceci économise énormément de temps et permet aux utilisateurs de rassembler seulement l'information dont ils ont besoin - et rien de plus.

Les informations des sources ouvertes en combinaison avec la traduction automatique par machine – un problème résolu ?

Le Basel Open Intelligence est aussi conçu afin de surmonter d'autres défis pratiques auxquels font face des enquêteurs, des analystes et des agents responsables de la conformité conduisant des vérifications des références ou des mesures de vigilance en utilisant des sources ouvertes.

Le langage est un défi de taille. Jusqu'à récemment, le problème était l'obstacle de la langue - que les analystes ne comprennent pas les informations fournies en langues étrangères. La traduction prend du temps et a un coût certain, surtout si vous ne savez pas si la source est pertinente ou pas. Les outils de traduction automatique comme DeepL et Google Translate ont

rendu plus rapide et plus économique la compréhension des informations des sources ouvertes en d'autres langues. Maintenant un nouveau problème se présente : les moteurs de recherche affichent surtout des résultats dans la langue de l'utilisateur. Il est donc probable que les analystes qui dépendent des moteurs de recherche habituels ne trouveront même pas l'informations cruciales disponibles dans d'autres langues, ni auront-ils l'occasion de l'analyser.

Afin de résoudre ce problème, le Basel Open Intelligence mène des recherches aux multiples langages. Il offre aux utilisateurs l'option de traduire automatiquement les articles dans leur langue afin d'en faciliter l'analyse.

Un avertissement

Quand il s'agit d'un renseignement d'origine sources ouvertes, les outils sont là pour aider les humains, pas pour les remplacer. Le langage est un parfait exemple : les recherches plurilingues et la traduction automatique sont des ressources précieuses, mais ne peuvent remplacer la compréhension humaine du contexte et des connotations culturelles ou politiques. Un mot comme « paramilitaire » ou même le mot « corruption » peuvent avoir des sens très différents dans le contexte des pays comme la Colombie, l'Irlande ou la Chine.

Les humains ont une meilleure vision de l'information manquante, pas seulement l'information trouvée grâce aux sources ouvertes. Même si beaucoup de registres d'entreprise et de bénéficiaires effectifs sont disponibles en ligne, d'autres manquent de transparence.

Surtout, l'informations des sources ouvertes sont une bête puissante qui a besoin d'être apprivoisée. La capacité d'extraire les informations pertinentes d'une masse d'information disponible sur Internet et ailleurs et de la transformer en quelque chose d'utile...ça, c'est le pouvoir. Voilà le renseignement.

Publié en anglais le 9 juin 2020

Tous nos guides rapides sont disponibles sur learn.baselgovernance.org

ISSN 2673-5229

Ce travail est sous licence Creative Commons Attribution-Non commercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0).

