



A recolha, tratamento e análise de fontes abertas de informação



Manuel Medina
Analista de Informação

Nas últimas décadas, as iniciativas de anticorrupção, transparência e liberdade de informação aumentaram significativamente a importância da recolha, tratamento e análise de fontes abertas de informação (intelligence), quer para o sector público, quer para o privado.

Neste guia rápido, o Analista de Informação, Manuel Medina, explica o que significa a recolha, tratamento e análise de fontes abertas de informação e explora algumas das questões complexas relacionadas com este tema.

A OSINT (Open Source Intelligence) e o seu enquadramento no ciclo da informação

A recolha, tratamento e análise de fontes abertas de informação, ou OSINT, sigla através da qual este serviço é conhecido na comunidade de informação, realiza a recolha, tratamento e análise sistemáticos de fontes abertas de informação, ou seja, da informação que é acessível ao público,

em geral, num registo comercial, como por exemplo artigos dos órgãos de informação, relatórios de pesquisa, ou registos de empresas.

Conforme expliquei no meu [guia rápido anterior sobre a recolha, tratamento e análise de informação](#), é importante considerar esta actividade não como um objecto estático, mas como um processo. O resultado deste processo pode ser um produto estratégico ou operacional. Em qualquer dos casos, é concebido para ajudar os decisores a ter um conhecimento profundo de um assunto ou ameaça específica, a fim de poderem tomar a decisão acertada para atingir objectivos concretos. Se esta actividade for considerada um processo, então as fontes abertas de informação representam um dos possíveis elementos envolvidos nesse processo.

As fontes abertas de informação por si só podem ser um instrumento poderoso para apoiar a tomada de decisão com conhecimento de causa. Mas, dependendo do contexto e da necessidade, os analistas também podem associar as informações de fontes abertas e as confidenciais no ciclo de informação.

Por exemplo, um informador ou denunciante pode dar uma dica de que uma determinada pessoa tem interesses empresariais não revelados e suspeitos. Podem usar-se fontes abertas de informação sobre registos de empresas, sites e artigos de órgãos de informação para tentar corroborar estes factos.

Poderá a OSINT servir de apoio à aplicação da lei?

Sim, na verdade, é essencial. Em muitos países, as notícias que circulam nos órgãos de informação, especialmente escritas por jornalistas de investigação, podem desencadear investigações em caso de suspeita de crimes financeiros. Em alguns casos, os indivíduos traem-se a si próprios, involuntariamente, através de informações que disponibilizam publicamente nas redes sociais. Isto aconteceu no caso da [família Ambuila](#) na Colômbia, cuja filha colocou fotografias suas online num Lamborghini e usando artigos de luxo.

De igual modo, como descreve o meu colega Tom Walugembe, o [primeiro caso de branqueamento de capitais no Uganda](#) começou com um vídeo bastante divulgado no WhatsApp, que mostrava os criminosos nos seus apartamentos, posando ao lado de maços de notas de dólares americanos. Um acto muito pouco inteligente.

Qual a origem da OSINT?

Temos tendência a associar o termo “fontes abertas” às informações digitais, mas também podem ser físicas. Antigamente, os analistas e investigadores costumavam ir a bibliotecas e repartições públicas para obter registos públicos e artigos. Certamente já viram filmes em que a polícia folheia pilhas de jornais velhos em arquivos empoeirados.

Hoje em dia, muitas dessas informações estão disponíveis na internet e

os analistas passam mais tempo sentados à secretária, mas ainda há no mundo muita informação de fontes abertas impressas em papel, especialmente em países com sistemas informáticos menos desenvolvidos.

Note que fácil acesso não significa que seja gratuito. Por exemplo, para se ter acesso a muitos jornais e artigos de órgãos de informação é necessário ser-se assinante pagante. Ter de pagar por informações publicamente disponíveis é uma questão duvidosa e, na verdade, uma dor de cabeça para as equipas de intelligence que têm poucos recursos.

Contudo, desde que o custo não seja exorbitante a ponto de ser proibitivo e que não haja outras restrições de acesso a essa informação, estes dados são classificados, de forma geral, como sendo de fontes abertas.

A importância da OSINT para os sectores público e privado

A recolha, tratamento e análise de fontes abertas de informação é de grande importância para a aplicação da lei. Gera pistas, corrobora informações recebidas de outras fontes e serve de suporte a provas utilizadas nos processos judiciais.

Também é extremamente útil para as empresas e instituições financeiras aquando da realização de avaliações de risco e diligências devidas aos clientes, funcionários e terceiros. As fontes abertas de informação podem ser usadas para identificar sistemas de fraude ou contrafacção, ou para investigar as delações feitas por um denunciante. De forma geral, as informações estratégicas empresariais também são baseadas em dados de mercado provenientes de fontes abertas.

Informação significa poder? Actualmente, já não

A rápida expansão da Internet e das ferramentas de pesquisa nas últimas duas décadas fez com que o público tivesse fácil acesso a uma enorme quantidade de informação. Agora bastam alguns cliques para nos levar a arquivos de jornais que remontam a vários anos, registos de decisões judiciais, ou bancos de dados estatísticos descarregáveis.

No entanto, paradoxalmente, ter acesso rápido e fácil a uma enorme quantidade de dados significa que o velho ditado “informação é poder” já não é assim tão verdadeiro.

Em primeiro lugar, há a questão das notícias falsas, isto é, artigos deliberadamente enganosos, ou publicações nas redes sociais sem qualquer fundamento. As notícias falsas que se tornam virais lançam achas para a fogueira (falsa). Os analistas têm de levar a cabo procedimentos morosos para verificar a fiabilidade das fontes abertas de informação, tanto durante a fase de recolha inicial, como durante as fases de integração e tratamento.

Em segundo lugar, demasiada informação pode “infoxicar-nos” facilmente e

atrapalhar o processo de tomada de decisão. Ter a capacidade de recolher dados relevantes e de extrair padrões úteis representa poder, não apenas informação.

Como combater a “infoxicação”

O combate à “infoxicação” é um problema grave. Se já passou horas completamente absorvido e perdido na internet enquanto pesquisava um assunto, saiba que não está sozinho. As ferramentas que ajudam a acelerar o processo de pesquisa, a extrair informações relevantes e a fazer as associações necessárias para se obter uma perspectiva global da questão, são realmente valiosas nestes casos.

A nossa própria ferramenta Basel Open Intelligence foi concebida para fazer exactamente isso, isto é, com um clique realiza uma busca automática do nome de um indivíduo, ou de uma organização, em conjugação com mais de 200 palavras-chave sobre crimes financeiros, acções judiciais, ou outras infracções penais, bem como listas de palavras-chave personalizadas. Também pesquisa a Web oculta, seleccionando quaisquer referências nas listas de sanções e pessoas politicamente expostas.

De forma vantajosa, é possível pesquisar, em simultâneo, vários pseudónimos e diferentes variações do nome. É incrível e ao mesmo tempo frustrante a quantidade de transliteraões e variações diferentes de nomes, especialmente no que diz respeito a línguas com diferentes alfabetos ou estruturas complexas dos nomes. Alguns indivíduos usam pseudónimos para esconder as suas verdadeiras identidades.

O algoritmo usa o processamento de linguagem natural para extrair entidades e revelar pessoas, empresas, localizações e cargos de responsabilidade. Os documentos encontrados durante a pesquisa são enumerados, juntamente com o texto principal extraído do site, excluindo conteúdos irrelevantes como anúncios, menus ou avisos de cookies. As palavras-chave são destacadas para facilitar a leitura. O utilizador pode filtrar e classificar rapidamente os documentos de acordo com a sua relevância.

Tudo isto poupa imenso tempo e permite aos utilizadores recolherem apenas as informações de que precisam.

Fontes abertas de informação e tradução automática – problema resolvido?

O Basel Open Intelligence também foi concebido para ajudar a superar outros desafios práticos enfrentados pelos investigadores, analistas e responsáveis pela conformidade, que realizam verificações dos antecedentes, ou diligências devidas através de fontes abertas de informação.

A língua é um grande problema. Até recentemente, o problema foi a barreira linguística, já que os analistas não entendiam as informações que chegavam em línguas que não falavam. A tradução leva tempo e é

dispendiosa, principalmente quando se desconhece se a fonte é relevante ou não.

As ferramentas de tradução automática como a DeepL e o Google Translate tornaram possível uma compreensão mais rápida e menos dispendiosa das fontes abertas de informação noutras línguas. Mas agora enfrentamos um novo problema: os motores de pesquisa normais procuram, em primeiro lugar, resultados na língua do utilizador. O mais provável é que os analistas que dependem de motores de pesquisa normais nem sequer encontrem informação potencialmente decisiva noutras línguas, quanto mais conseguirem analisá-la.

Para ajudar a resolver isto, o Basel Open Intelligence efectua pesquisas em múltiplas línguas. Também dá aos utilizadores a opção de traduzir automaticamente os artigos para a sua própria língua para facilitar a análise.

Uma palavra de prudência

Quando se trata de fontes abertas de informação, existem ferramentas para ajudar os humanos e não para os substituir. A língua é um bom exemplo: a pesquisa multilingue e a tradução automática são extremamente valiosas, mas não substituem a compreensão humana do contexto e das conotações culturais ou políticas. Uma palavra como “paramilitar”, ou mesmo “corrupção”, pode ter significados muito diferentes no contexto da Colômbia, Irlanda ou China.

Os humanos também têm uma noção mais exacta da informação que está em falta, e não apenas da informação que podem encontrar através de fontes abertas. Embora muitos registos de empresas, ou dos proprietários efectivos, estejam online e acessíveis, há outras que carecem de transparência.

Acima de tudo, as fontes abertas de informação são uma ferramenta muito poderosa, mas é preciso saber dominá-la. A capacidade de extrair informação relevante da enorme quantidade de informação disponível publicamente na Internet e noutros locais, e de a transformar em algo útil... a isso sim, chama-se poder. Isso é *intelligence*.

Publicado em inglês a 9 de Junho de 2020

Todos os nossos guias rápidos estão disponíveis em learn.baselgovernance.org

ISSN 2673-5229

Esta obra está licenciada sob uma Licença Creative Commons Atribuição-Não-comercial-NoDerivs 4.0 Internacional (CC BY-NC-ND 4.0).



Basel Institute on Governance
Steinenring 60
4051 Basel, Switzerland

+41 61 205 55 11
info@baselgovernance.org
baselgovernance.org

 @BaselInstitute
 Basel Institute

The Basel Institute on Governance
is an Associated Institute of the
University of Basel.

