

Classificação por Ranqueamento de Acesso: Análise Web em Ferramentas de Inteligência de Fontes Abertas

João Rafael Gonçalves Evangelista (Universidade Nove de Julho) jrafa1607@gmail.com

Dacyr Dante de Oliveira Gatto (Universidade Nove de Julho) dacyr.gatto@uni9.pro.br

Renato José Sassi (Universidade Nove de Julho) sassi@uni9.pro.br

Resumo:

Encontram-se na internet inúmeras ferramentas que realizam buscas em fontes públicas, por exemplo, em mídias sociais. A busca, coleta e análise de informações de fontes públicas para propósitos inteligentes, é denominada inteligência de fontes abertas ou como é conhecida OSINT (*Open Source Intelligence*). Com diversas ferramentas OSINT disponíveis na internet, é difícil saber qual é a mais utilizada em sua área de atuação. Um possível método para entender quais são as ferramentas OSINT mais utilizadas é por meio da análise web, técnica que extrai indicadores do comportamento dos usuários nas páginas na internet. Com base nesse cenário, o objetivo deste trabalho foi realizar uma análise web em ferramentas OSINT para classificá-las pelo total de acessos recebidos utilizando o indicador de ranqueamento de acesso. A metodologia de pesquisa adotada foi a descritiva com abordagem quantitativa. Quanto aos procedimentos técnicos, este trabalho é classificado como experimental, por verificar se a aplicação da análise web em ferramentas OSINT é capaz de produzir indicadores para realizar uma classificação. Os resultados evidenciaram que os indicadores proporcionados pela análise web foram capazes de classificar as ferramentas OSINT pelo ranqueamento de acesso.

Palavras chave: OSINT, Inteligência de Fontes Abertas, Análise Web, Ranqueamento de Acesso.

Classification by Access Ranking: Web Analytics in Open Source Intelligence Tools

Abstract:

There are many tools on the internet that search in public sources, for example, in social media. The search, collection and analysis of information from public sources for intelligent purposes is called Open Source Intelligence or as OSINT is known. With many OSINT tools available on the internet, it is difficult to know which one is most used in your area of expertise. A possible method to understand which OSINT tools are most used is through web analytics, a technique that extracts indicators of users' behavior in web pages. Based on this scenario, the objective of this work was to perform a web analytics in the OSINT tools to classify them by the total number of accesses received using the access ranking indicator. The research methodology adopted was descriptive with a quantitative approach. As for the technical procedures, this work is classified as experimental, by verifying if the application of web analysis in OSINT tools is able to produce indicators to perform a classification. The results showed that the indicators provided by the web analysis were able to classify the OSINT tools by the access ranking.

Key-words: OSINT, Open Source Intelligence, Web Analytics, Access Ranking.

1. Introdução

As pessoas estão cada vez mais divulgando informações na internet, especialmente em mídias sociais. Práticas como a publicação de listas de funcionários em sites organizacionais permitem que pessoas com más intenções identifiquem facilmente os funcionários de uma empresa entre milhões de usuários de mídias sociais (EDWARDS et al., 2017).

Medkova (2018) aponta que embora a publicação de dados em mídias sociais traga vantagens significativas, como a coleta de informações valiosas para pesquisadores em diferentes áreas, como marketing, psicologia social e segurança da informação. Também causam problemas de segurança da informação, por exemplo na preservação de privacidade, já que os dados de redes sociais contêm muitas vezes, informações confidenciais e privadas.

Encontram-se disponibilizadas na internet inúmeras ferramentas que realizam a busca de informações em fontes públicas. Estas ferramentas fazem buscas nas mais variadas fontes, como: blocos de código, mídias sociais, páginas acadêmicas e sites de relacionamento.

A capacidade das ferramentas de busca de entender, selecionar, organizar e usar estas informações é crucial para transformar as informações vastas, mas dispersas, em inteligência humana e computacional, apoiando a aprendizagem e o desempenho humano, bem como o desenvolvimento de sistemas e serviços inteligentes (CHEN; DÉCARY, 2018).

Segundo Glassman e Kang (2012), o termo utilizado para descrever a aquisição de dados de fontes publicamente disponíveis, bem como as técnicas utilizadas para adquirir esses dados é denominada inteligência de fontes abertas ou como é conhecida OSINT (*Open Source Intelligence*).

Chauhan e Panda (2015) apontam alguns tipos de informação que podem ser extraídas de fontes públicas, como: Publicações acadêmicas, dados de mídias sociais e dados de órgãos públicos. A extensão na qual as informações podem ser coletadas é grande e o campo onde as informações se encontram é amplo (WATTERS; LAYTON, 2016).

Com inúmeras ferramentas disponíveis na internet que fazem busca em fontes públicas, é difícil saber quais são as mais utilizadas. Uma possível maneira de classificar e descobrir quais ferramentas OSINT são as mais utilizadas é por meio da análise web.

Čegan e Filip (2017) definem a análise web como um campo científico que engloba diversas subdisciplinas, como medição do tráfego da web, coleta de dados grandes, análise de desempenho da web, mineração de dados corporativos e estratégias de visualização de dados.

A análise web proporciona indicadores capazes de analisar e classificar páginas na internet, por exemplo: Total de visitas mensais que a página recebeu, tipo do dispositivo que efetuou o acesso (mobile ou desktop), o tempo médio de duração de cada visita ou mesmo o número médio de páginas visitadas dentro daquele domínio.

2. Objetivo

O objetivo deste trabalho foi realizar uma análise web em ferramentas de inteligência de fontes abertas (OSINT) para classificá-las pelo total de acessos recebidos utilizando o indicador de ranqueamento de acesso, e assim, descobrir quais são as ferramentas de inteligência de fontes abertas mais utilizadas.

3. Fundamentação Teórica

3.1. Inteligência de Fontes Abertas

Segundo Koops, Heopman e Leenes (2013), inteligência de fontes abertas (OSINT) envolve o processo de coleta, análise e uso de dados de fontes públicas para propósitos inteligentes. Já para Howells e Ertugan (2017), OSINT é uma forma de gerenciamento de coleta de inteligência que localiza, seleciona e extrai informações de fontes abertas, como *Twitter* e *Facebook*, e por fim, analisa as informações para produzir inteligência pela forma que podem ser usadas.

Segundo a metodologia de testes de segurança da informação *PTES Technical Guideline*, OSINT, no mais simples dos termos, é localizar e analisar fontes de informação disponíveis publicamente (abertas). Na área de segurança da informação, este processo de coleta de informações tem o objetivo de produzir informações atuais e relevantes que sejam valiosas para um invasor ou um concorrente.

Quick e Choo (2018) apontam alguns exemplos de fontes públicas onde a OSINT pode atuar, como: mídias globais, blogs na internet, relatórios governamentais, imagens de satélite, trabalhos acadêmicos, *Wikipedia*, *Youtube* e *Facebook*, além de uma série de outras informações por meio da internet e de outros recursos de mídia.

Pela riqueza, variedade e importância da informação que circula dentro do conceito de OSINT, as organizações (especialmente as mais sensíveis às temáticas da segurança da informação) começam a reconhecê-lo como um ponto crucial da atividade de monitorização de segurança (RODRIGUES, 2015).

Clarke (2015) define a informação descoberta por OSINT como informação disponível publicamente que qualquer pessoa obteria legalmente por solicitação, compra ou observação que atenda um requisito de inteligência. Esta informação é o ingrediente inicial para o processo de análise pelo qual a inteligência ou o conhecimento é criado.

Hribar, Podbregar e Ivanuša (2014), Akhgar, Bayerl e Sampson (2017), Chauhan e Panda (2015), Quick e Choo (2018) apresentam outras disciplinas de coleta de inteligência, em que OSINT pode atuar diretamente com cada uma. As disciplinas de inteligência são apresentadas a seguir na tabela 1.

Disciplinas de Inteligência	Descrição
COMINT	Inteligência de Comunicação
CULTINT	Inteligência Cultural
DFINT	Inteligência de Forense Digital
ELINT	Inteligência de Eletrônica
GEOINT	Inteligência Geoespacial
HUMINT	Inteligência Humana
IMINT	Inteligência de Imagens
MARKINT	Inteligência de Mercado
MASINT	Inteligência de Medição e Assinatura
SIGINT	Inteligência de Sinais
SOCMINT	Inteligência de Mídias Sociais
TECHINT	Inteligência Técnica
TELINT	Inteligência de Telemetria

Fonte: Adaptado de Akhgar, Bayerl e Sampson (2017); Chauhan e Panda (2015); Hribar, Podbregar e Ivanuša (2014); Quick e Choo (2018).

Tabela 1 – Disciplinas de Inteligência

Glassman e Kang (2012) sugerem que OSINT não seja um tipo completamente novo de inteligência, mas em geral, surgiu na resolução de problemas humanos somente durante tipos específicos de atividades direcionadas por objetos. Sua prática é vista geralmente em termos positivos, particularmente como um método de coleta de dados convencional que não viola os direitos humanos (HRIBAR; PODBREGAR; IVANUŠA, 2014).

A prática de coleta de dados vem sendo discutida desde 1941, quando um esforço para monitorar transmissões de rádio alemãs e japonesas foi lançado com a criação do *Foreign Broadcast Monitoring Service*, uma organização que mais tarde se transformou em Centro de Fontes Públicas (CLARKE, 2015).

A partir da criação do Centro de Fontes Públicas até os dias atuais, surgiram inúmeras ferramentas e técnicas de coleta de informações de fontes abertas que automatizam a busca e análise direcionadas a um determinado objeto de inspeção. Lee e Shon (2016) por exemplo, abordam a prática de ferramentas OSINT como *Google*, *Shodan*, *Sensys*, *theHarvester*, *Z-map* e *Carrot2*, para encontrar vulnerabilidades.

3.2. Análise Web

Segundo Salini, Malavolta e Rossi (2016), a análise web trata-se do uso de ferramentas e *softwares* que coletam dados obtidos do estudo do comportamento dos usuários enquanto eles navegam na internet. Estes dados podem ser obtidos por exemplo, rastreando os cliques do mouse e solicitações de informações dos visitantes do website.

Beasley (2013) define a análise web como a maneira de aprender como os usuários interagem com sites e aplicativos móveis, registrando automaticamente aspectos do comportamento dos usuários e, em seguida, combinando e transformando o comportamento em dados que podem ser analisados.

A análise web é amplamente usada para saber informações sobre onde o tráfego da web está surgindo, em que tipos de produtos os visitantes estão interessados, que tipos de palavras-chave os usuários estão digitando nos mecanismos de pesquisa para acessar um site, etc. (RESHMA; RAJENDRAN, 2017).

Um exemplo de aplicação da análise web é a análise do conteúdo gerado pelo usuário em uma mídia social, como opiniões sobre produtos. A organização pode usar essas opiniões como *feedback* em relação ao seu produto para melhorá-lo, enquanto o cliente pode usar as mesmas opiniões das massas nestas mídias sociais em relação a um produto para decidir se ele é digno de comprar o produto ou não (RATHAN et al., 2017).

Outro exemplo de aplicação abordado por Järvinen e Karjaluo (2015), se trata do uso da análise web para realizar a medição de desempenho no marketing digital. Já Fagan (2014), apresenta a análise web para obter e avaliar os indicadores de desempenho gerados por estudantes universitários dentro de uma biblioteca.

Choi et al. (2018) apresentam o uso da análise web por meio da ferramenta *SimilarWeb* para desenvolver uma categorização de páginas web. Enquanto Zhang (2014) utilizou-se da mesma ferramenta para explorar o interesse e o uso do site PhET (Página de simulações interativas em ciências e matemática) em uma universidade.

4. Metodologia

4.1 Tipo de Pesquisa

Esta pesquisa é classificada como descritiva, pois inspeciona a aplicação da análise web em ferramentas OSINT. A pesquisa descritiva é definida por Gil (2008) como pesquisa que têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis. Suas características mais significativas estão na utilização de técnicas padronizadas de coleta de dados.

Quanto a natureza dos dados desenvolvidos nesta pesquisa, este trabalho apresenta uma abordagem quantitativa, visto que a pesquisa trabalha com a análise dos dados levantados e processados em um software de análise web, com o propósito de se obter indicadores para aplicar uma classificação.

Em relação aos procedimentos técnicos, esta pesquisa é classificada como pesquisa experimental por investigar se a aplicação de análise web é capaz de classificar ferramentas OSINT. A pesquisa experimental consiste em determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo, definir as formas de controle e observação dos efeitos que a variável produz no objeto (GIL, 2008).

Sobre a fundamentação teórica deste trabalho, esta pesquisa realizou um levantamento bibliográfico por meio das palavras chaves: “OSINT”, “Open Source Intelligence” e “Web Analytics” nas bases de periódicos: IEEE Digital Library, Scielo, Scopus, ScienceDirect, EmeraldInsight, Portal Capes e ProQuest.

4.2 Condução dos Experimentos Computacionais

Os experimentos computacionais foram divididos em seis fases, descritas a seguir:

- a) Fase A: Busca de ferramentas OSINT: Procurou-se por conjuntos de ferramentas OSINT para retirar a amostra de ferramentas OSINT necessária para o desenvolvimento da pesquisa. Para isso, utilizou-se as seguintes ferramentas de busca: “Carrot2” <<http://search.carrot2.org/stable/search>> e Google <<https://www.google.com/>>.
- b) Fase B: Busca de ferramentas de análise web: Procurou-se nos conjuntos de ferramentas OSINT encontrados na Fase A por ferramentas de análise web para realizar a classificação do experimento. Após definir os requisitos, foram selecionadas duas ferramentas de análise web: *Similarweb* <<https://www.similarweb.com/>> e *Crunchbase* <<https://www.crunchbase.com/>>.
- c) Fase C: Avaliação de ferramentas de análise web: Avaliou-se as ferramentas de análise web descobertas na Fase B, e assim, pôde-se identificar quais destas ferramentas é capaz de produzir os indicadores necessários para realizar uma classificação. Para isso, executou-se uma análise web na rede social *LinkedIn* <<http://www.linkedin.com/>>, com as ferramentas descobertas na fase B.
- d) Fase D: Definição da amostra de ferramentas OSINT: Definiu-se a amostra das ferramentas OSINT que será avaliada neste trabalho. Para isso, consultou-se na ferramenta de análise web *SimilarWeb* definida na Fase C, quais são os 100 domínios na internet mais acessados. Após isso, pegou-se os 5 domínios mais bem classificados e verificou-se nos conjuntos de ferramentas OSINT como estes domínios estão agrupados.
- e) Fase E: Extração de ferramentas OSINT: Retirou-se a amostra de ferramentas OSINT definidas na etapa D e as inseriu em uma planilha no *Microsoft Excel 2016* junto com um ID e seu link de acesso.

f) Fase F: Avaliação e Classificação das ferramentas OSINT: Avaliou-se as ferramentas OSINT extraídas na Fase E por meio de uma análise web realizada pela ferramenta *SimilarWeb*. Feita a análise web, retirou-se os indicadores desejados e os inseriu na planilha criada na fase E. Após obter estes indicadores, criou-se um atributo denominado “Ranqueamento de Acesso” para assim, classificar as ferramentas OSINT pelo total de visitas recebidas.

Apresenta-se na figura 1 as fases dos experimentos computacionais.

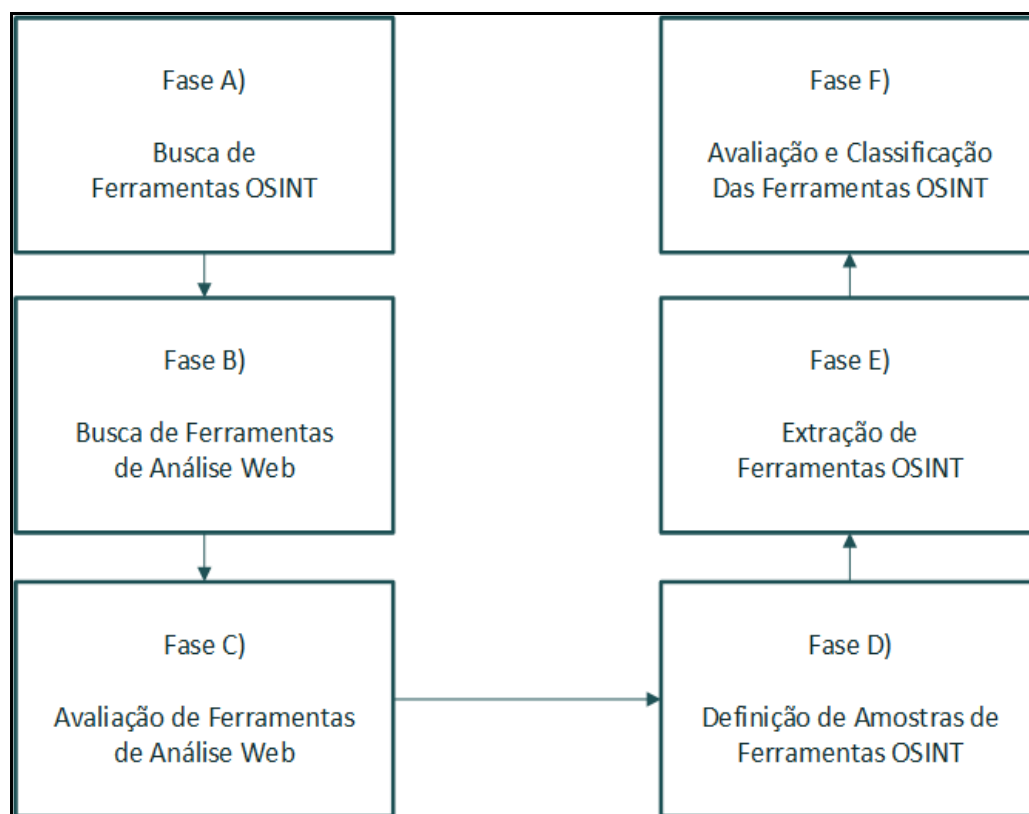


Figura 1 – Fases dos experimentos computacionais

5. Apresentação e Discussão dos Resultados

Nesta seção, são apresentados os resultados obtidos dos experimentos computacionais, como: Conjuntos de ferramentas OSINT utilizadas, seleção e avaliação de ferramentas de análise web, definição e extração da amostra das ferramentas OSINT, execução da análise web em ferramentas OSINT e por fim, a classificação das ferramentas OSINT pelo total de visitas recebidas. Os resultados dos experimentos computacionais são descritos a seguir.

a) Fase A: Busca de ferramentas OSINT: Procurou-se por conjuntos de ferramentas OSINT com as ferramentas de busca: “Carrot2” e “Google” com as palavras chaves: “OSINT Framework” e “Plataforma OSINT”.

Para critério de seleção dos conjuntos de ferramentas OSINT, verificou-se quais dos conjuntos encontrados trariam a maior quantidade de ferramentas OSINT agrupadas por categorias de atuação e quais deles são mencionados em livros ou publicações acadêmicas. Na tabela 2 apresentada a seguir, são descritos os conjuntos de ferramentas OSINT encontrados, juntamente com seu tipo e link de acesso.

Conjunto de Ferramentas OSINT	Tipo de Conjunto	Link de Acesso
OsintFramework	Página Web	http://osintframework.com/
I – Intelligence Osint Tools and Resources Handbook	Livro Digital (Pdf)	https://www.i-intelligence.eu/wp-content/uploads/2016/11/2016_November_Open-Source-Intelligence-Tools-and-Resources-Handbook.pdf

Fonte: Autores (2018).

Tabela 2 – Conjuntos de Ferramentas OSINT

b) Fase B: Busca de ferramentas de análise web: Na tabela 3 apresentada a seguir, são demonstradas as ferramentas de análise web encontradas nos conjuntos de ferramentas OSINT definidos na Fase A. Foi utilizado para critério de seleção quais ferramentas de análise web poderiam ser executadas de forma livre por um período mínimo de 30 dias.

Ferramentas de Análise Web	Link de Acesso
SimilarWeb	https://www.similarweb.com/
Crunchbase	https://www.crunchbase.com

Fonte: Autores (2018).

Tabela 3 – Ferramentas de Análise Web

c) Fase C: Avaliação de ferramentas de análise web: Realizou-se uma avaliação nas ferramentas de análise web encontradas na fase B. Para isso, executou-se uma análise web na rede social *LinkedIn* <<http://www.linkedin.com>> com as ferramentas descobertas na fase B: *Crunchbase* e *SimilarWeb*. A ferramenta *Crunchbase* disponibilizou muito mais informações qualitativas do que quantitativas, sendo incapaz de descobrir indicadores sobre o uso dos usuários com a rede social testada (*LinkedIn*). A análise web por meio ferramenta *Crunchbase* disponibilizou valores como: Nome dos fundadores, endereços de e-mail de alguns colaboradores, investidores, links para mídias sociais, preço atual da organização, nome da equipe da organização, dentre outras.

Diferente da ferramenta *Crunchbase*, a ferramenta *SimilarWeb* disponibilizou mais informações quantitativas sobre a interação do usuário com a rede social *LinkedIn*, como: Rank global, total de visitas recebidas, média de visitas mensais, tempo médio de cada visita e taxa média de rejeição. Sendo assim, escolheu-se a ferramenta *SimilarWeb* para executar a análise web neste trabalho. Os indicadores escolhidos para realizar a classificação de ferramentas de inteligência de fontes abertas pelo total de visitas recebidas foram: Rank global e total de visitas recebidas no mês de julho de 2018 (data mais recente na execução deste trabalho).

d) Fase D: Definição da amostra de ferramentas OSINT: Retirou-se da ferramenta *SimilarWeb* o rank global com os 100 domínios mais acessadas entre os meses de maio de 2018 até julho de 2018 em todo o mundo, e a partir desta informação, verificou-se nos conjuntos de ferramentas OSINT como as ferramentas estão agrupadas. O rank global informado pela ferramenta *SimilarWeb* apresentado na figura 2 a seguir.

	Domain	Traffic Share	Change	Rank	Monthly Visits	Unique Visitors	Desktop vs Mobile
1	 google.com	9.80%	↓ -0.22%	#1	43.71B	2.226B	40.54% 59.4
2	 youtube.com	5.56%	↑ 0.98%	#2	24.79B	1.895B	62.22% 37.7
3	 facebook.com	5.08%	↓ -2.17%	#3	22.64B	1.623B	43.74% 56.2
4	 baidu.com	2.45%	↑ 7.05%	#4	10.92B	527.7M	8.25% 91.7
5	 wikipedia.org	1.22%	↑ 0.48%	#10	5.446B	1.168B	43.18% 56.8

Figura 2 – Domínios mais acessados entre maio e julho de 2018

O rank global informado pela ferramenta *SimilarWeb* mostra que as ferramentas classificadas como: Ferramentas de Busca nos conjuntos definidos na fase A, ocupam duas posições (1^o *Google* e 4^o *Baidu*), dos cinco domínios mais acessados.

Sendo assim, consultou-se nos conjuntos de ferramentas OSINT definidos na fase A, quais são as ferramentas OSINT classificadas como ferramentas de busca e categorias semelhantes, estão disponíveis. As ferramentas OSINT encontradas estão classificadas como: Ferramentas de Busca Geral, Ferramentas de Metabúscua e Ferramentas de Busca por Agrupamentos.

e) Fase E: Extração de ferramentas OSINT: Extraíu-se as ferramentas OSINT da amostra definida na fase D. A tabela 4 apresenta a seguir as ferramentas extraídas. Nesta tabela, se encontram o ID da ferramenta, o nome da ferramenta e seu link de acesso.

ID	Ferramentas de Análise Web	Link de Acesso
1	Advangle	http://advangle.com/
2	All-io.net	http://all-io.net/
3	Alltheinternet	https://www.alltheinternet.com/
4	Aol	http://search.aol.com
5	Ask	http://www.ask.com
6	Baidu	http://www.baidu.com/
7	Bing	https://www.bing.com/
8	Bing Vs Google	http://bvsg.org/
9	Biznar	https://biznar.com
10	Carrot2	http://search.carrot2.org/stable/search
11	Dogpile	http://www.dogpile.com/
12	Dothop	http://dothop.com
13	DuckDuckGo	https://duckduckgo.com/
14	Etools	https://www.etools.ch/
15	Exalead	https://www.exalead.com/search/
16	Excite	http://www.excite.com
17	Factbites	http://www.factbites.com
18	Faganfinder	http://www.faganfinder.com/
19	Gigablast	http://gigablast.com
20	Goodsearch	http://www.goodsearch.com
21	Goofram	http://www.goofram.com/
22	Google	http://www.google.com
23	Info	http://www.info.com
24	Infospace	http://www.infospace.com
25	Instya	http://www.instya.com/#/web/
26	Izito	https://www.izito.com/
27	Kartoo	http://www.kartoo.com/
28	Lycos	http://www.lycos.com/
29	Metabear	http://www.metabear.com/
30	Mojeek	https://www.mojeek.com/

31	Mywebsearch	https://hp.mywebsearch.com
32	Nextaris	https://www.nextaris.com/
33	Oscobo	https://oscobo.co.uk
34	Parseek	http://www.parseek.com/
35	Qwant	https://www.qwant.com/
36	Search	http://www.search.com
37	Searx	https://searx.me/
38	Startpage	https://www.startpage.com/
39	Surfcanyon	http://www.surfcanyon.com/
40	Swisscows	https://swisscows.com/
41	Teoma	http://www.teoma.com/
42	Trovando	http://www.trovando.it/
43	Webcrawler	http://www.webcrawler.com/
44	Yahoo	http://yahoo.com/
45	Yandex	https://www.yandex.com/
46	Zapmeta	https://www.zapmeta.com

Fonte: Autores (2018).

Tabela 4 – Ferramentas OSINT

f) Fase F: Avaliação e classificação de ferramentas OSINT: Após extrair as ferramentas OSINT, pegou-se o link de acesso das mesmas e os inseriu na ferramenta *SimilarWeb*. Para a data de referência da avaliação, utilizou-se o mês de julho de 2018 por se tratar da data mais recente disponibilizada no momento da pesquisa. Feito a análise web, selecionou-se os indicadores necessários para realizar a classificação.

Os indicadores selecionados foram: O total de acessos que a ferramenta recebeu no mês de julho de 2018 e o *rank* global. Inicialmente, classificou-se as ferramentas OSINT pelo total de visitas recebidas. Como algumas destas ferramentas OSINT receberam valor de “ < 5000 ” no total de visitas, foi necessário apoiar-se em outro indicador. Por isso, apoiou-se nos dados do rank global informados pela ferramenta *SimilarWeb* para assim, realizar a classificação.

Por fim, criou-se um novo atributo denominado “Rank Acesso” contendo a classificação das ferramentas OSINT pelo total de visitas recebidas. A tabela 5 apresenta a seguir as ferramentas OSINT classificadas pelo ranqueamento de acesso.

ID	Ferramentas de Análise Web	Link de Acesso	Rank Global	Total de Visitas Recebidas	Rank Acesso
22	Google	http://www.google.com	1	43.800.000.000	1
6	Baidu	http://www.baidu.com/	4	10.200.000.000	2
44	Yahoo	http://yahoo.com/	6	4.753.000.000	3
7	Bing	https://www.bing.com/	39	1.226.000.000	4
13	DuckDuckGo	https://duckduckgo.com/	94	478.300.000	5
5	Ask	http://www.ask.com	319	188.900.000	6
38	Startpage	https://www.startpage.com/	445	120.600.000	7
35	Qwant	https://www.qwant.com/	1058	61.040.000	8
45	Yandex	https://www.yandex.com/	659	52.590.000	9
4	Aol	http://search.aol.com	N/A	23.580.000	10
24	Infospace	http://www.infospace.com	4309	20.120.000	11
11	Dogpile	http://www.dogpile.com/	9794	7.201.000	12
23	Info	http://www.info.com	13427	6.016.000	13
43	Webcrawler	http://www.webcrawler.com/	14308	5.442.000	14
16	Excite	http://www.excite.com	10452	5.435.000	15
26	Izito	https://www.izito.com/	14964	4.846.000	16
46	Zapmeta	https://www.zapmeta.com/	21898	3.310.000	17
36	Search	http://www.search.com	30130	2.904.000	18

41	Teoma	http://www.teoma.com/	33172	2.455.000	19
28	Lycos	http://www.lycos.com/	41102	1.243.000	20
37	Searx	https://searx.me/	56963	1.229.000	21
34	Parseek	http://www.parseek.com/	76716	999.756	22
3	Alltheinternet	https://www.alltheinternet.com/	100540	651.603	23
31	Mywebsearch	https://hp.mywebsearch.com	N/A	323.415	24
40	Swisscows	https://swisscows.com/	219350	240.044	25
15	Exalead	https://www.exalead.com/search/	236043	215.223	26
19	Gigablast	http://gigablast.com	241874	143.929	27
14	Etools	https://www.etools.ch/	324536	129.144	28
10	Carrot2	http://search.carrot2.org/stable/search	N/A	52.760	29
20	Goodsearch	http://www.goodsearch.com	1298027	38.704	30
30	Mojeek	https://www.mojeek.com/	1076564	35.040	31
12	Dothop	http://dothop.com	956706	30.539	32
18	Faganfinder	http://www.faganfinder.com/	1405382	26.100	33
27	Kartoo	http://www.kartoo.com/	1741402	19.475	34
33	Oscobo	https://oscobo.co.uk	3092479	15219	35
2	All-io.net	http://all-io.net/	2537405	13.739	36
39	Surfcanyon	http://www.surfcanyon.com/	2720944	10.500	37
17	Factbites	http://www.factbites.com	2311097	9.890	38
8	Bing Vs Google	http://bvsg.org/	3437335	7.047	39
9	Biznar	https://biznar.com	6194392	< 5000	40
42	Trovando	http://www.trovando.it/	7182221	< 5000	41
21	Goofram	http://www.goofram.com/	8912852	< 5000	42
1	Advangle	http://advangle.com/	9729414	< 5000	43
29	Metabear	http://www.metabear.com/	12136910	< 5000	44
25	Instya	http://www.instya.com/#/web/	15532695	< 5000	45
32	Nextaris	https://www.nextaris.com/	38426380	< 5000	46

Fonte: Autores (2018).

Tabela 5 – Classificação de ferramentas OSINT pelo ranqueamento de acesso.

6. Considerações Finais.

Neste trabalho abordou-se a classificação por ranqueamento de acesso de ferramentas de inteligência de fontes abertas (OSINT) por meio de análise web com o objetivo de evidenciar aquelas que são as mais acessadas.

A aplicação da análise web por meio da plataforma *SimilarWeb* gerou importantes indicadores, dos quais, foram utilizados o “Total de Visitas recebidas” e “Rank Global”. Indicadores estes, que foram capazes de classificar e ranquear as ferramentas OSINT pelo ranqueamento de acesso.

Assim, pode-se perceber quais ferramentas OSINT destacam-se segundo suas características e atuação pela quantidade de visitas recebidas. Como sugestão para trabalhos futuros, será interessante dar continuidade na avaliação das ferramentas OSINT, pois incorporando as demais categorias e não somente as categorias apresentadas neste trabalho, torna-se possível o desenvolvimento de um *framework* ou plataforma OSINT voltada para marketing ou mesmo, segurança da informação.

Referências

- AKHGAR, BABAK; BAYERL, P. SASKIA; SAMPSON, FRASER (Ed.).** *Open Source Intelligence Investigation: From Strategy to Implementation*. Springer, 2017.
- BEASLEY, MICHAEL.** *Practical web analytics for user experience: How analytics can help you understand your users*. 1st Edition. USA. Newnes, 2013.
- CARROT2.** Carrot2 Clustering Engine. Disponível em: < <http://search.carrot2.org/stable/search> >. Acesso em 6. Maio. 2018.
- ČEGAN, LUKÁŠ; FILIP, PETR.** Webalyt: Open web analytics platform. In: *Radioelektronika (RADIOELEKTRONIKA)*, 2017 27th International Conference. IEEE, p. 1-5, 2017.
- CHAUHAN, SUDHANSHU; PANDA, NUTAN KUMAR.** *Hacking Web Intelligence: Open Source Intelligence and Web Reconnaissance Concepts and Techniques*. 1st Edition. USA. Syngress, 2015.
- CHEN, MEI; DÉCARY, MICHEL.** A Cognitive-Based Semantic Approach to Deep Content Analysis in Search Engines. In: *Semantic Computing (ICSC)*, 2018 IEEE 12th International Conference on. IEEE, 2018. p. 131-139. 2018.
- CHOI, DAEJIN; HAN, JINYOUNG; CHUN, SELIN; RAPPOS, EFSTRATIOS; ROBERT, STEPHAN; KNOW, TED TAEKYOUNG.** Bit.ly/practice: Uncovering content publishing and sharing through URL shortening services. *Telematics and Informatics*, v.35, Issue 5, p. 1310-1323, 2018.
- CLARKE, CRAIG. S.** Open Source Intelligence. An oxymoron or real intelligence? *Marine Corps Gazette – Professional Journal of U.S. Marines*. v.99, Issue 8, p. 22, 2015.
- CRUNCHBASE.** Ferramenta de Análise Web. Disponível em: < <https://www.crunchbase.com/> >. Acesso em 25. Maio. 2018.
- EDWARDS, MATTHEW; LARSON, ROBERT; GREEN, BENJAMIM; RASHID, AWAI; BARON, ALISTAIR.** Panning for gold: Automatically analysing online social engineering attack surfaces. *Computers & Security*, v. 69, p. 18-34, 2017.
- FAGAN, JODY CONDIT.** The suitability of web analytics key performance indicators in the academic library environment. *The Journal of Academic Librarianship*, v. 40, n. 1, p. 25-34, 2014.
- GIL, ANTÔNIO CARLOS.** *Métodos e técnicas de pesquisa social*. 6. ed. Editora Atlas SA, 2008.
- GLASSMAN, MICHAEL; KANG, MIN JU.** Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior*, v. 28, n. 2, p. 673-682, 2012.
- GOOGLE.** Google Search Engine. Disponível em: < <https://www.google.com.br/> >. Acesso em 6. Maio. 2018.
- HOWELLS, KAREN; ERTUGAN, AHMET.** Applying fuzzy logic for sentiment analysis of social media network data in marketing. *Procedia Computer Science*, v. 120, p. 664-670. 2017
- HRIBAR, GAŠPER; PODBREGAR, IZTOK; IVANUŠA, TEODORA.** OSINT: a “grey zone”? *International Journal of Intelligence and CounterIntelligence*, v. 27, n. 3, p. 529-549, 2014.
- I – INTELLIGENCE.** *Osint Tools and Resources Handbook*. Disponível em: < >. Acesso em 10. Maio. 2018.
- JÄRVINEN, JOEL; KARJALUOTO, HEIKKI.** The use of Web analytics for digital marketing performance measurement. *Industrial Marketing Management*, v. 50, p. 117-127, 2015.
- KOOPS, BERT-JAAP; HOEPMAN, JAAP-HENK; LEENES, RONALD.** Open-source intelligence and privacy by design. *Computer Law & Security Review*, v. 29, n. 6, p. 676-688, 2013.
- LEE, SEOKCHEOL; SHON, TAESHIK.** Open source intelligence base cyber threat inspection framework for critical infrastructures. In: *Future Technologies Conference (FTC)*. IEEE, p. 1030-1033. 2016.
- LINKEDIN.** Rede Social Profissional. Disponível em: < <https://br.linkedin.com/> >. Acesso em 27. Maio. 2018.
- MEDKOVA, JANA.** Composition attack against social network data. *Computers & Security*, v. 74, p. 115-129, 2018.
- OSINTFRAMEWORK.** Framework focado na coleta de informações utilizando ferramentas ou recursos livres. Disponível em < <http://osintframework.com/> >. Acesso em 16. Setembro. 2018

QUICK, DARREN; CHOO, KIM-KWANG RAYMOND. *Digital forensic intelligence: Data subsets and Open Source Intelligence (DFINT+ OSINT): A timely and cohesive mix.* *Future Generation Computer Systems*, v. 78, p. 558-567, 2018.

PENETRATION TESTING EXECUTION STANDARD – TECHNICAL GUIDELINES. Disponível em < http://www.pentest-standard.org/index.php/PTES_Technical_Guidelines>. Acesso em 16. Setembro. 2018

RATHAN, M.; VISHWANATH, R HULIPALLED; MURUGESHWARI, P.; SUSHMITHA, H.M. *Every post matters: A survey on applications of sentiment analysis in social media.* In: *Smart Technologies For Smart Nation (SmartTechCon), International Conference On. IEEE*, 2017. p. 709-714. 2017.

RESHMA, K.; RAJENDRAN, V. V. *An enhanced approach for querying integrated web analytics ontology using Quepy.* In: *Intelligent Computing and Control (I2C2), 2017 International Conference on. IEEE*, p. 1-6. 2017.

RODRIGUES, BERNARDO DE SIMAS GASPAS. *Open-source intelligence em sistemas SIEM.* Universidade de Lisboa – Faculdade de Ciências. Dissertação de Mestrado. 2015.

SALINI, ANDREA; MALAVOLTA, IVANO; ROSSI, FABRIZIO. *Leveraging web analytics for automatically generating mobile navigation models.* In: *Mobile Services (MS), 2016 IEEE International Conference on. IEEE*, 2016. p. 103-110. 2016.

SIMILARWEB. *Plataforma de Análise Web.* Disponível em: < <https://www.similarweb.com/pro> >. Acesso em 25. Maio. 2018.

WATTERS, P. A.; LAYTON, R. *Automating Open Source Intelligence: Algorithms for OSINT.* USA. Syngress, 2016.

ZHANG, MEILAN. *Who are interested in online science simulations? Tracking a trend of digital divide in Internet use.* *Computers & Education*, v. 76, p. 205-214, 2014.