

O uso de Informação Geográfica Voluntária e Operadores de Busca Avançados na web como ferramentas metodológicas para a construção de bancos de dados de eventos meteorológicos-oceanográficos Intensos/extremos no litoral paulista

XIV ENCOGERCO



Encontro Nacional de Gerenciamento Costeiro

Silva, V.D.M.¹, Linhares, P.S.², Souza, C.R.G.³

¹ Departamento de Geografia - Universidade de São Paulo

² Programa de Pós-Graduação de Geografia Física – Universidade de São Paulo

³ Instituto de Pesquisas Ambientais

INTRODUÇÃO

O uso da Web acelerou a formação de uma cultura voltada à produção e consumo de informações que vêm sendo aproveitadas por pesquisas científicas no campo social e em novos métodos de coleta e processamento de dados (CAMPAGNA, 2016; FRAGOSO; RECUERO; AMARAL, 2011). Ademais, com o estabelecimento de telefonia móvel, GPS, pontos de wifi, bluetooth e smartphones ocorre a coexistência de um ambiente de informação digital online com o espaço físico, gerando produtos colaborativos, de acordo com o conceito de Informação Geográfica Voluntária (GOODCHILD, 2007). Assim, considerou-se a ubiquidade e o compartilhamento de conteúdo por usuários Web para coleta de informações de eventos meteorológicos e oceanográficos no litoral paulista, além dos prejuízos ambientais, pois os usuários são atores locais que registram os fenômenos ocorridos em sua comunidade e/ou em área de convivência.

METODOLOGIA

A forma de coleta dos dados foi o somatório de Operadores de Busca Avançados (RUSSEL, 2021) com Termos de Busca, escolhidos com critérios geográficos e conceituais, no motor de busca Pesquisa Google, que posteriormente foram tabulados por meio do Google Forms. Os Operadores Avançados de Busca trabalham com a linguagem HTML, gerando comandos que localizam termos no título da página (<title>), corpo do texto (<body>) ou no URL. Posteriormente, os dados foram estruturados (WICKHAM, 2014) para programas de computador, uso em Sistemas de Informações Geográficas (SIGs) e aplicação em visualizadores de dados. A pesquisa compreende o período do ano 2000 a 2020. Conforme a tabela 01, propôs-se inicialmente uma lista abrangente de termos de busca com relação muito próxima aos eventos de interesse e a escolha deu-se por pesquisa prévia e discussões com a equipe de pesquisa do Projeto Ressacas. Em seguida, os termos foram categorizados em quatro grupos que buscam ser mutuamente excludentes (CARLOMAGNO; CAETANO DA ROCHA, 2016), sem ambiguidades e estritos, apesar de haver dentro de cada um dos grupos 1 e 2 uma quantidade significativa de sinônimos.

Grupo 1: lista os eventos climáticos, oceanográficos, locais e objetos geográficos ou geomorfológicos típicos de regiões litorâneas;

Grupos 2: Consequências e atividades causadas por excesso de chuva, ressacas e altas de maré.

Grupo 3: lista dos 16 municípios que compõem a Baixada Santista e o Litoral Norte;

Grupo 4: Operadores Avançados de Pesquisa

GEOGRÁFICOS	CLIMÁTICOS	OCEANOGRÁFICOS
beira-mar	chuva	mar
	garoa	maré alta
costa	tempestade	maré-cheia
	temporal	oceano
praia	ventania	ressaca
	vento	
CONSEQUÊNCIAS		OPERADOR
agitação	enxurrada	allintext:
agitado	erosão	allintitle:
alagamento	estrago	allinurl:
ameaça	impacto	site:facebook.com
avaria	inundação	site:twitter.com
balanço	movimentado	site:youtube.com
cheia	perda	
dano	perigo	
destruição	prejuízo	
enchente	risco	

Tabela 01. Lista da categorização dos termos de busca. O grupo 3 foi omitido, pois os resultados consideraram apenas um município

RESULTADOS E DISCUSSÃO

Os termos de busca simples foram combinados de forma a evitar ambiguidades e que garantissem a informação espacial necessária, totalizando 27.648 termos de busca compostos, para todo o litoral paulista. Diante do elevado número de termos de busca, considerou-se como teste as publicações realizadas no município de Santos, pois o termo “santos” tem elevada quantidade de sentidos, como time de futebol, sobrenome e símbolos religiosos, e também é o município com o maior contingente de veículos jornalísticos (PROJOR, 2021). Sem qualquer método de filtragem, o mesmo termo retorna centenas de milhões de resultados. A metodologia reduziu para 3.648 inserções de termos de busca e a avaliação se deu da seguinte maneira (ANDRADE, 2018): operadores de buscas retornam documentos com links aproveitáveis para o banco de dados (SINAL) ou não aproveitáveis (RUIÍDO); quando o motor de busca não retornou resultados, foi classificado como SEM RESPOSTAS. Os Termos de Busca Avançados, no geral, retornaram cerca de 5% de links aproveitáveis, 35,1% com resultados e sem aproveitamento e 59,9% operadores que não retornaram resultados. Quando analisamos as respostas para os termos simples nos grupos 1 e 2, pode-se ver quais palavras são mais produtivas para o criação banco de dados, de acordo com a Tabela 02. Da mesma forma, os melhores Termos de Busca Avançadas aparecem na Tabela 03, onde evidencia-se que termos localizados no corpo do texto <body> são mais eficientes e dependendo da localização do termo na página HTML, pode não haver respostas.

GRUPO 1				GRUPO 2			
SINAL		SEM RESPOSTA		SINAL		SEM RESPOSTA	
chuva	13,59%	costa	8,65%	agitação	11,96%	avaria	5,91%
maré alta	10,87%	litorâneo	8,65%	inundação	10,33%	movimentado	5,72%
beira-mar	10,33%	beira-mar	7,65%	agitado	9,78%	prejuízo	5,59%
ressaca	9,78%	ventania	7,33%	erosão	9,78%	perigo	5,54%
temporal	9,24%	ressaca	7,28%	destruição	8,15%	perda	5,49%

Tabela 02. Lista de termos buscas simples de acordo com respostas, não considerando ruídos

SINAL	RUIÍDO	SEM RESPOSTA
allintext:santos agitação costa	allintext:santos estrago vento	site:twitter.com santos balanço ressaca
allintext:santos destruição costa	allintext:santos destruição temporal	site:facebook.com santos avaria beira-mar
allintext:santos agitação ressaca	site:facebook.com santos prejuízo tempestade	site:twitter.com santos prejuízo garoa
allintext:santos consequência chuva	allintext: santos inundação ressaca	site:facebook.com santos risco costa
allintext:santos agitado costa	allintext:santos dano região costeira	allintitle:santos movimentado maré alta
site:youtube.com santos inundação maré alta	site:youtube.com santos avaria vento	allinurl:santos impacto chuva
site:youtube.com santos estrago beira-mar	site:facebook.com santos avaria chuva	allintitle:santos ameaça beira-mar
allintext:santos destruição ressaca	site:facebook.com santos prejuízo oceano	allintitle:santos agitação oceano
allintext:santos estrago beira-mar	allintext:santos perigo tempestade	allinurl:santos agitado costa
allintext:santos agitação maré alta	site:facebook.com santos destruição vento	allinurl:santos balanço maré-cheia

Tabela 03. Lista de Termos de Busca Avançados conforme às respostas

CONCLUSÃO

Em síntese, o uso de operadores de busca avançados e a seleção de melhores termos de busca auxiliam na pesquisa científica. Contudo, se faz necessária uma breve compreensão de como o conteúdo em sites noticiosos e mídias sociais digitais são publicados. Em áreas que envolvam a produção de informações por meio de Sistemas de Informação Geográficas e, consequentemente, sua utilização para o escopo de gestão ambiental, há aproveitamento para o trabalho do pesquisador. Os dados fornecidos voluntariamente por usuários da Web são atuais e podem abranger uma grande área territorial e um extenso período temporal.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, Sidgley C. De; CASTRO DEGRASSI, Lívia; RESTREPO-ESTRADA, Camilo; DELBEM, Alexandre C. B.; PORTO DE ALBUQUERQUE, João. Does keyword noise change over space and time? A case study of social media messages. In: XIX GEOINFO 2018, Campina Grande. Anais [...]. Campina Grande p. 116-121. Disponível em: <http://mtc-m16c.sid.inpe.br/col/sid.inpe.br/mtc-m16c/2018/12/27.18.33/doc/p11.pdf>.
- CAMPAGNA, Michele. Social Media Geographic Information: Why social is special when it goes spatial? In: European Handbook of Crowdsourced Geographic Information. [s.l.] : Ubiquity Press, 2016. p. 45-54. Disponível em: http://www.ubiquitypress.com/site/chapters/10_5334/bax_d/. Acesso em 02/03/2021.
- CASTELLS, Manuel. A galáxia da internet. Reflexões sobre a internet, os negócios e a sociedade. 1ª ed. [s.l.] : Zahar, 2013.
- FRAGOSO, Suelly; RECUERO, Raquel; AMARAL, Adriana. Métodos de Pesquisa para Internet. Porto Alegre: Sulina, 2011. v. 1
- GOODCHILD, Michael F. Citizens as sensors: The world of volunteered geography. GeoJournal. [S. l.], v. 69, n. 4, p. 211-221, 2007.
- PROJOR - Instituto para o Desenvolvimento do Jornalismo. Os desertos de Notícia no Brasil. Atlas da Notícia. Disponível em: <https://www.atlas.jor.br/desertos-de-noticia/>. Acesso em: 25/03/2021
- RUSSEL, Daniel M. Advanced Search Operators, 2019. Disponível em: <https://sites.google.com/site/dmrussell/>. Acesso em 02/03/2021.
- WICKHAM, Hadley. Tidy data. Journal of Statistical Software, [S. l.], v. 59, n. 10, p. 1-23, 2014.