

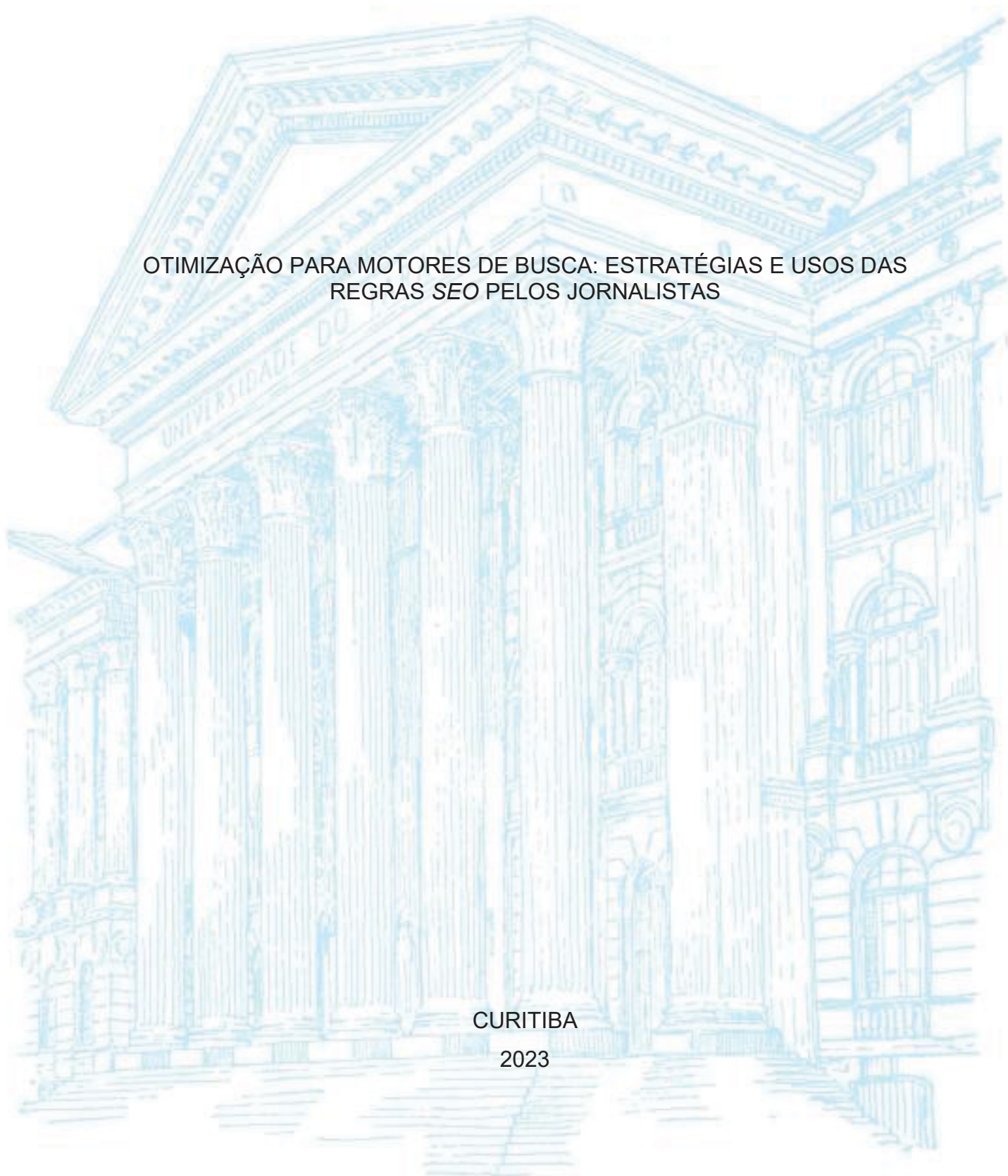
UNIVERSIDADE FEDERAL DO PARANÁ

IANA CRISTINA LIEBEL

OTIMIZAÇÃO PARA MOTORES DE BUSCA: ESTRATÉGIAS E USOS DAS
REGRAS SEO PELOS JORNALISTAS

CURITIBA

2023



GIANA CRISTINA LIEBEL

OTIMIZAÇÃO PARA MOTORES DE BUSCA: ESTRATÉGIAS E USOS DAS
REGRAS SEO PELOS JORNALISTAS

Dissertação apresentada ao Programa de Pós-graduação em Comunicação, Setor de Artes, Comunicação e Design, da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestra em Comunicação.

Linha: Comunicação e Cultura.

Orientador: Prof. Dr. Rodrigo Eduardo Botelho-Francisco

CURITIBA

2023

UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS
BIBLIOTECA DE ARTES COMUNICAÇÃO E DESIGN - CABRAL

L716 Liebel, Giana Cristina
 Otimização para motores de busca: estratégias e usos das regras
 SEO pelos jornalistas. / Giana Cristina Liebel. – 2023.
 1 recurso online : PDF

 Orientador: Prof. Dr. Rodrigo Eduardo Botelho-Francisco
 Dissertação (mestrado) – Universidade Federal do Paraná, Setor de
 Artes, Comunicação e Design, Programa de Pós-graduação em
 Comunicação.

 Inclui referências.

1. Comunicação. 2. Jornalistas. 3. Ciberjornalismo. 4. *Google*. I.
Botelho-Francisco, Rodrigo Eduardo. II. Universidade Federal do
Paraná. Setor de Artes Comunicação e Design. Programa de Pós-
graduação em Comunicação. III. Título.

CDD: 302.2



MINISTÉRIO DA EDUCAÇÃO
SETOR DE ARTES
COMUNICAÇÃO E DESIGN
UNIVERSIDADE FEDERAL DO
PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-
GRADUAÇÃO PROGRAMA DE PÓS-
GRADUAÇÃO COMUNICAÇÃO -
40001016071P8

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação COMUNICAÇÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **GIANA CRISTINA LIEBEL** intitulada: **Otimização para motores de busca: estratégias e usos das regras SEO pelos jornalistas**, sob orientação do Prof. Dr. RODRIGO EDUARDO BOTELHO FRANCISCO, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 11 de Abril de 2023.

Assinatura Eletrônica 17/04/2023 15:04:36.0

RODRIGO EDUARDO BOTELHO FRANCISCO
Presidente da Banca Examinadora

Assinatura Eletrônica 17/04/2023 15:24:42.0

MÁRCIO CARNEIRO DOS SANTOS
Avaliador Externo (UNIVERSIDADE FEDERAL DO MARANHÃO)

Assinatura Eletrônica 17/04/2023 15:24:10.0

MYRIAN REGINA DEL VECCHIO DE LIMA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Rua Bom Jesus, 650 - CURITIBA - Paraná - Brasil
CEP 80035-010 - Tel: (41) 3313-2063 - E-mail: ppgcom@ufpr.br
Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 276644
Para autenticar este documento/assinatura, acesse <https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jspe>
insira o código 276644

RESUMO

Com inúmeros textos sendo inseridos na web por minuto, cada vez mais é necessário entender melhor em como conseguir estar entre as dez primeiras posições de uma ferramenta de pesquisa da Internet. Com isso, esta dissertação teve como objetivo geral responder à seguinte questão: o Jornalista se utiliza de quais estratégias para ranquear no buscador Google? Para tanto, vimos a necessidade de explicar o que é o *Search Engine Optimization (SEO)*, bem como de alguns dos conceitos que norteiam o ciberespaço, o qual o Jornalismo se insere nos tempos atuais, além da própria plataforma Google. Realizamos também um estado da arte em que o *SEO* e jornalismo foram as buscas principais, porém, pelo baixo volume apresentado, acrescentamos a busca por uma regra específica do *SEO*, o *Tag Title*. Do ponto de vista metodológico, utilizamos um estudo de caso a partir do portal de notícias do programa Globo Esporte (GE) e uma entrevista com o coordenador desse site de Jornalismo Esportivo, para trazer luz sobre a utilização dos resultados encontrados nesta dissertação na visão dos jornalistas, como por exemplo, a de que a *Tag Title* não é mais utilizada pelo algoritmo do Google tanto quanto a metadescrição.

Palavras-Chaves: Comunicação; Jornalista; Ciberjornalismo; *SEO*; Google; *Tag-Title*.

ABSTRACT

With countless texts being added to the Web per minute, it is increasingly necessary to understand how to be among the top ten positions of an Internet search engine. With this in mind, this dissertation aimed to answer the question: Do journalists employ strategies to be among the top results in the Google search engine? To this end, we explained what Search Engine Optimization (SEO) is and all the concepts that guide cyberspace—which is now a backdrop of journalism practices—as well as the Google platform itself. We also conducted a state-of-the-art review using "SEO" and "journalism" as the primary search terms, and later (given the initially low number of results) added another term to cover a more specific SEO rule, "Tag Title". Furthermore, our methodology included a case study on the sports news website Globo Esporte (GE) and an interview with its coordinator to shed light on the use of the results found in this dissertation, e.g., how Tag Title is no longer used by Google algorithm as much as the meta description.

Keywords: Communication; Journalist; Cyberjournalism; SEO; Google; Tag Title

SUMÁRIO

1 INTRODUÇÃO.....	2
1.1 PROBLEMÁTICA.....	3
1.2 OBJETIVOS.....	6
1.3 JUSTIFICATIVA.....	6
1.4 ESTRUTURA DA DISSERTAÇÃO.....	7
2 O JORNALISMO NO CIBERESPAÇO.....	9
2.1 A PLATAFORMA GOOGLE.....	20
2.2 <i>SEO</i> e <i>TAG TITLE</i>	27
2.3 MODELO DE ANÁLISE.....	30
3. PROCEDIMENTOS METODOLÓGICOS.....	33
3.1 ESTADO DA ARTE.....	33
3.2 PESQUISA DOCUMENTAL GOOGLE.....	46
3.3 ESTUDO DE CASO.....	46
4. RESULTADOS E ANÁLISES.....	51
5. CONSIDERAÇÕES FINAIS.....	54
6. REFERÊNCIAS.....	56
7. APÊNDICE.....	64

1 INTRODUÇÃO

A informação em um clique. Em 2023 parece algo inconcebível uma pessoa não se atentar a pesquisar sobre alguma notícia, informação na Internet. Passados mais de 30 anos da sua implementação, ocorrida em 1990 (SCHWINGEL, 2012) pelo físico britânico Tim Berners-Lee, a *world wide web* tem sido utilizada sem discriminação. Aqui não vamos nos atentar a fatores econômicos, políticos ou territoriais, mas sim, às questões de inserção de informação. A *www*, como é mais conhecida, designa um sistema de documentos em hipermídia que são interligados e executados na Internet, por meio de links. Os documentos podem estar na forma de vídeos, sons, hipertextos e imagens. Um grande marco para o Jornalismo foi quando os primeiros jornais digitais passaram a estar disponíveis na íntegra e de forma gratuita - primeiro nos Estados Unidos, em 1993, e no Brasil, em 1995 (SCHWINGEL, 2012):

David Carlson (2003, pp 49-50) afirma: o primeiro sítio jornalístico na web foi lançado em novembro de 1993, na Faculdade de Jornalismo e Comunicação da Universidade da Flórida, e, em 19 de janeiro de 1994, o *Palo Alto Weekly*, na Califórnia, tornou-se o primeiro jornal a publicar regularmente na *web*. (SCHWINGEL, 2012, p. 9).

Assim como na televisão, que trouxe muito da literatura impressa em suas transmissões, o Jornalismo Digital também nasceu atrelado aos conceitos do impresso, visto que sua formulação estava vinculada ao processo de informatização das redações. Porém, Mielniczuk e Palacios (2001) atentam-se para as constantes alterações que estavam ocorrendo nesse meio, e definitivamente, distanciando-se da metáfora do impresso.

Tal distanciamento do digital para o impresso se dá, principalmente, pelas características que encontramos no ciberjornalismo: hipertexto, multimídia, interatividade, memória, entre outros que serão citados mais adiante. Se em 2022 qualquer pessoa pode inserir um texto (será citado sempre como texto todo e qualquer formato que se pode carregar na Internet, como foto, vídeo, gráfico), temos aqui, uma influência da tecnologia crescente em se tratando de gerenciamento de dados. Indexação, armazenamento e recuperação viraram ações normais no cotidiano quando a questão é a busca e inserção de informações (HOLANDA; BRAZ, 2012). Essa influência é exponencial principalmente em uma sociedade onde

a explosão documental atingiu seu ápice e os usuários sozinhos não conseguem obter as informações desejadas (MAIMONE; SILVEIRA, 2007).

Segundo Holanda e Braz (2012), a troca de informações na Internet atingiu proporções que não foram ainda alcançadas por outro meio de comunicação, uma vez que, se tornou um mecanismo de disseminação de dados, independentemente de suas localizações geográficas. Além do mais, seu crescimento é exponencial, e, a cada dia, o número de páginas disponíveis aumenta, chegando à estimativa do surgimento de 320 milhões de endereços por semana.

Com esse contexto, se na década de 1970 as redes de comunicação brasileiras monopolizaram o direito à produção e publicação de informações - herança política dos governos militares que passou a estar atrelada a interesses econômicos, segundo Schwingel (2012) - as mídias online representam possibilidades para que o Jornalismo exerça sua função pública de informar o cidadão para que seja livre em sua forma de pensar e capaz de se autogovernar, conforme afirmam Kovach e Rosentiel (2004). E é no início de 2000 que o ciberespaço começa a ser visto pelos seus usuários como, além de uma fonte para consulta, uma nova mídia, um novo suporte, um novo meio de comunicação (SCHWINGEL, 2012), assim como o rádio, o jornal impresso e a televisão, com o ciberjornalista como profissional atrelado ao meio.

Nesse contexto, a intenção desta dissertação foi levantar os principais conceitos que norteiam o ciberjornalismo, adequando-se às características do ciberespaço, e, principalmente para que as produções jornalísticas apareçam na primeira página dos mecanismos de busca, utilizando-se, aqui, das regras *SEO*, e como essa técnica norteia a estratégia do jornalista e os referenciais acadêmicos. A discussão teórica terá como base os conceitos de cibercultura (LÉVY, 1999), de Jornalismo online (PALACIOS, 1999; FERRARI, 2003; CAVAVILHAS, 2014) e ciberjornalismo (SCHWINGEL, 2012; PALACIOS, 2011, BARBOSA, 2004, 2013).

1.1 PROBLEMÁTICA

Pensar a localização de conteúdos na Internet envolve, desde 1998, considerar a utilização do Google, que ao longo de pouco mais de 20 anos se consolidou como a principal ferramenta de busca na rede mundial de computadores. Segundo Nikolovska (2019), ele domina a preferência entre as plataformas de pesquisa,

sendo priorizado por 80% das pessoas ao considerarem resultados orgânicos ao invés dos pagos.

Com um elevado nível de disseminação de informações, de diferentes tipos de narradores sobre todos os acontecimentos do mundo, o Google, em 2021, teve mais de sete bilhões de *pageviews* por dia, sendo realizadas mais de 87,8 bilhões de buscas por mês, segundo o portal Tecmundo, especializado em informações tecnológicas.

Por conta da infinidade de conteúdos, o principal mecanismo de busca criou suas próprias regras para organizar seu sistema, cujo *slogan* aparece em uma de suas principais páginas: “Nossa missão é organizar as informações do mundo para que sejam universalmente acessíveis e úteis para todos”¹.

Desde 2011 o Google vem atualizando seus algoritmos com o intuito de melhorar essa organização dos conteúdos inseridos na Internet, bem como ajustar a melhor entrega para a pesquisa. Em 2011, por exemplo, ele penalizava os sites que abusavam da inserção de palavras-chave em seus conteúdos. Já em 2015 o fator de expansão de uso da plataforma por meio de telefones celulares começou a fazer efeito na busca de melhor ranqueamento. O sistema chega em 2019 com a incorporação de um sistema de inteligência artificial, que segundo o próprio Google, é especializado no processamento da linguagem natural para compreender como o ser humano faz buscas. Tais técnicas padronizadas são mais conhecidas como *SEO* (*Search Engine Optimization*), ou seja, Otimização do Mecanismo de Busca. Como exemplo, uma das regras inseridas no *SEO*, é a técnica da *TagTitle*.

A *TagTitle*, antes de ser lida pelos seres humanos no ciberespaço, é analisada pelos algoritmos responsáveis por essa organização. Uma das regras formatadas pelo Google é a de que ela tenha até 60 caracteres (exatamente para ele ser totalmente exibido no *feed* de resultados). Limitações como essa não são novidades para o Jornalismo. O “Manual de Redação e Estilo do O Estado de S. Paulo” (2001), por exemplo, já citava que para o título, a utilização predominante de verbos deveria ser no infinitivo, sem utilização de ponto final, e indicava que se devia respeitar o limite de sinais para cada título, caso contrário, ele seria reduzido.

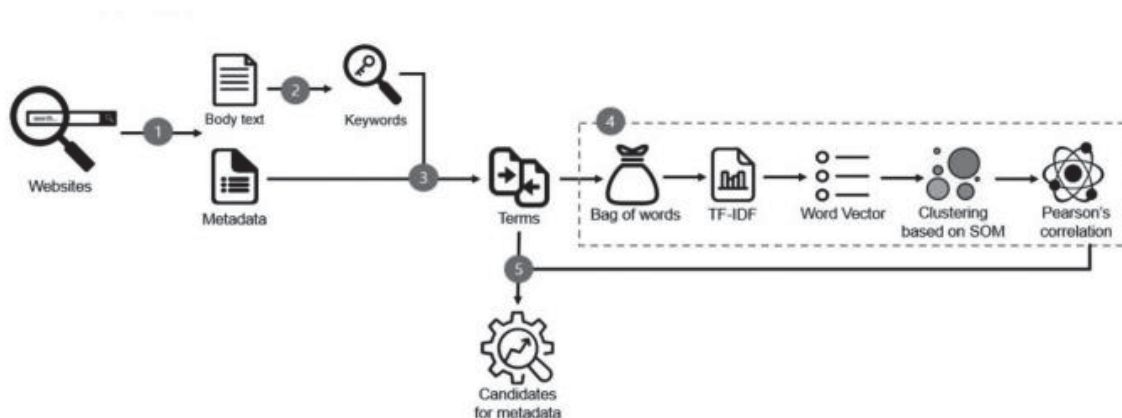
Por outro lado, em termos comunicacionais, são os títulos resumem a notícia (MELO, 2010). Para Burnett (1991, *apud* Bertolini, 2012) o título é o elemento mais

¹ Google - Sobre (about.google) - acessado em 26 de julho de 2022 <https://about.google/intl/pt-BR>.

importante da produção jornalística, pois, esse elemento é, muitas vezes, tudo que o leitor vai ler sobre o assunto ou é o fator que vai motivá-lo ou não a enfrentar a leitura do texto. Na Internet, isso é especialmente verdadeiro, porque o leitor só chega na notícia se clicar no título.

Antes de seguirmos para a pergunta principal, é importante situar o leitor ao que se refere à *SERP* (*Search Engine Results Page*), ou a Página de Resultados do Mecanismo de Busca. Abaixo temos uma imagem (AN; JUNG, 2021) para melhor exemplificar o que ela faz:

Imagem 1: Funcionamento da SERP



Segundo os autores, a partir de uma consulta do usuário, um algoritmo gera candidatos a metadados nos sites. A estrutura é a seguinte: (1) Análise de sites, (2) Extração de palavras-chave, (3) Combinação entre palavras-chave e metadados em termos, (4) Medição da relevância semântica de termos com base na *SERP*, (5) Computação da importância dos termos por *SERP* e relevância semântica. Ou seja, a *SERP* ajusta, por ordem de importância, os melhores resultados de acordo com cada pesquisa do interagente da Internet. A importância de estar nas primeiras posições do ranqueamento, além de financeiramente lucrativa para todas as empresas, não apenas as de notícias, é porque, como veremos mais a fundo nos próximos tópicos, 80% das pessoas prioriza a primeira página de resultados, deixando de seguir para as páginas seguintes.

Portanto, se, de um lado está o jornalista, com todas as regras *SEO* para desenvolver o texto, do outro, o interagente no ciberespaço, ao pesquisar sobre um assunto, depara-se no meio de um processo com um intermediador: o algoritmo. Neste contexto, pergunta-se: como é feita a escolha das melhores estratégias, pelo

profissional da escrita, para que seu texto seja ranqueado na primeira página do Google?

1.2 OBJETIVOS

Identificar quais são as principais estratégias que o jornalista pode utilizar para ranquear sua matéria na primeira página.

E os subsequentes objetivos específicos, foram:

1. Apresentar os conceitos que norteiam o *SEO* como um todo e, também, a regra específica *TagTitle*, analisando quais referenciais teóricos já existem e que englobam ambos;
2. Verificar e compreender a lógica de utilização da regra *SEO* e *TagTitle* em um caso específico;
3. Identificar as visões dos jornalistas sobre Cibercultura, técnicas de ranqueamento e definições de *SEO*.

1.3 JUSTIFICATIVA

Neste trabalho, temos a intenção de atualizar as questões já escritas sobre as regras *SEO*, podendo ser utilizadas como uma referência para acadêmicos e profissionais da Comunicação que tem a intenção de se aperfeiçoarem sobre as regras pré-estabelecidas pelo Google, além de aprimorar a cidadania digital, quando consideramos empoderar as pessoas de uma visão crítica sobre o funcionamento das tecnologias digitais. Economicamente, como já nos trouxe Myers (1998), a relação entre a pesquisa científica, a pesquisa corporativa e a exploração dos produtos comercializados em decorrência das descobertas provenientes dessa relação, atua como força disseminadora tanto da constante inovação, como na ampla oferta e comercialização de bens tecnológicos.

E por fim, outro propósito foi aumentar o conhecimento dessa jornalista de formação, especialista em marketing, que tantas vezes já se viu na frente de um sistema de produção de conteúdo para a Internet, como o WordPress tentando

imaginar qual seriam as melhores estratégias a serem utilizadas no texto para ser ranqueado pelo Google.

1.4 ESTRUTURA DA DISSERTAÇÃO

Na apresentação do nosso referencial teórico trazemos conceitos, características e a historicidade do ciberjornalismo, muitas vezes com um diálogo junto ao Jornalismo impresso. Utilizamos reflexões e conceitos de Barbosa (2008), Santos (2013), Ferraril (2003), Corrêa, Bertocchi (2012), Schwingel (2012), entre outros. Já quando falamos sobre a plataforma de busca Google, trazemos a própria empresa como referencial, buscando compreender um pouco mais sobre como a plataforma se comporta quando o assunto é o ranqueamento, além de alguns autores como Wise; Malseed (2007), Palacios, (2016), Holanda; Braz, (2012). Finalizando a etapa do referencial, passamos para o estudo sobre *SEO* e *TagTitle* e referenciamos Teixeira e Schiel (1997), Mercado (2002), entre outros autores encontrados por meio de pesquisa de estado da arte.

Nos procedimentos metodológicos apresentamos como foi feito nosso estado da arte, baseado em uma pesquisa dos artigos que referenciavam *SEO* Jornalismo e *TagTitle* em algumas bases de dados, além de um estudo de caso no site *Globo Esporte* (GE). A escolha desse portal se deu em decorrência de uma observação de 14 dias em cima da palavra “Copa do Mundo”. O termo foi encontrado por meio do *Google Trends*, uma plataforma que indica quais são os principais assuntos do momento. O portal GE foi o mais citado e ranqueado nesses dias de observação. O intuito desse estudo é tentar mostrar como a empresa GE entrega suas matérias aos interagentes do portal e se elas apresentam algum padrão, justificados pelas regras *SEO*. Alinhado com uma pesquisa documental feita com base na política de privacidade e nos termos de serviço do Google para enumerar as principais questões que a empresa apresenta aos usuários para que eles possam utilizar sua plataforma, verificando com isso, questões como usabilidade, interatividade, mercado e interface. Além disso, foi feita uma entrevista com o coordenador do portal *Globo Esporte*, da filial de Curitiba-PR, para verificar a visão do jornalista sobre *SEO*, *TagTitle*, Ciberjornalismo e a plataforma Google.

Nos resultados trouxemos algumas demandas a serem observadas, como por exemplo, a quantidade de artigos em português em muito menor número do que a

de inglês; algumas regras de *SEO* com peso maior do que outras; o alerta da utilização de uma plataforma de automatização do conteúdo trabalhada pela redação de portais, uma vez que as *tags* inseridas nas matérias, em 90% dos links observados, são iguais ao título (*TagTitle*) e iguais ao subtítulo (*metadescription*); a importância da *TagTitle* para os portais.

2 O JORNALISMO NO CIBERESPAÇO

Para que entendamos o espaço micro, é importante contextualizar o espaço macro. As regras *SEO* encontram-se no interior da cibercultura (LÈVY, 1999), em que se situa o ciberjornalismo, ou Jornalismo simplesmente, já que a prática, em geral, se apresenta no formato digital online com frequência cada vez maior e mais ampla.

Há muito que aplicativos, algoritmos e banco de dados são tidos como manifestações ciberculturais contemporâneas e alteram formas de sociabilidade (CORRÊA; BERTOCCHI, 2012. p. 3). Mas, como comenta Lèvy (1999, p 27) “o crescimento do ciberespaço não determina automaticamente o desenvolvimento da inteligência coletiva, apenas fornece a esta inteligência um ambiente propício”, ou seja, as informações estão ali, o ser interagente não entenderá apenas por ler as informações. Lèvy conceitua a inteligência coletiva como uma inteligência múltipla, compartilhada pela memória, distribuída por todo lugar, resultando em um aprendizado coletivo e na troca de conhecimentos. Ainda para o autor, o ciberespaço é onde a comunicação se interconecta, mundialmente, pelos computadores e suas memórias. A cada minuto que passa, novas pessoas passam a acessar a Internet, novos computadores são interconectados, novas informações são injetadas na rede. “Quanto mais o ciberespaço se amplia, mais ele se torna universal, e menos o mundo informacional se torna totalizável” (1999, p. 110). Comparando-se às mídias anteriores, que praticam uma comunicação mais unidirecional, quando os receptores estão isolados uns dos outros, como ouvindo uma estação de rádio ou vendo um canal da televisão, o ciberespaço encoraja a troca recíproca.

Nesse ambiente propício para um desenvolvimento de uma inteligência coletiva, a navegação do interagente é conduzida por um sistema de busca, baseado em um algoritmo – nesta dissertação, entenderemos algoritmo como um protocolo padrão de varredura do conteúdo dos dados disponíveis na rede (CORRÊA; BERTOCCHI, 2012. p. 6). Contudo, há de se notar que os aspectos comunicacionais, em um primeiro momento são feitos com um diálogo entre homem-máquina. E como sintetizam Molich e Nielsen (1990, p. 339), “a máquina não entende a linguagem humana, e esse papel é feito pela interface, quando um comando é dado e não é possível ser compreendido pela máquina, se denota a presença do ruído”.

Entendemos aqui um ruído como uma entrega não 100% condizente com o que se procura. Mas nosso foco aqui não será entender a parte máquina, e sim, a parte humana: a grande quantidade de informações injetadas na *web* trouxe o *information overload*², ou seja, “quando se torna insuportável a quantidade de dados disponíveis, o usuário recorre aos especialistas, delegando a eles a tarefa de organizar e dar sentido aos dados. No caso da informação noticiosa, espera-se que o jornalista execute essa tarefa” (CORRÊA; BERTOCCHI, 2012. p. 4).

Consideremos aqui, como um exemplo de parte do ciberespaço e para entendermos como funciona a quantidade de conteúdos existentes nele, a plataforma Google (falaremos mais a fundo sobre ela nos próximos capítulos):

Nessa empresa, o Google *Spider* ou *Googlebot* (algoritmos utilizados como rastreadores de informações) faz uma varredura nos dados tagueados e organizados no espaço de metadados propiciado pela estrutura da linguagem HTML de uma página; faz o mesmo sucessivamente para todas as páginas endereçadas na rede; organiza estas palavras-chave num índice de dados; quando um usuário alimenta a ferramenta com um termo de busca, o sistema processa tal lista de palavras-chave e rastreia os sites que contêm a mesma, apresentando o resultado por meio de um algoritmo – o *PageRank*, que os ordena pela frequência com que a palavra aparece na página, pelo tempo que a mesma está no ar e pelo número de outras páginas que remetem links para a mesma, entre outras alternativas (CORRÊA; BERTOCCHI, 2012. p. 6).

Nesse processo, com o contexto de que os humanos esperam que o comunicador organize as principais informações, cabe ao jornalista um domínio na construção de um conteúdo de página, além de passível de ser tagueada para essa varredura, dentro das regras de programação, mas também com um conteúdo visível para o buscador por meio de técnicas de *SEO* (*Search Engine Optimization*). Em suma, é um processo em que a ação comunicativa é determinada pelo sistema de busca (CORRÊA; BERTOCCHI, 2012. p. 7).

Antes de discorrer sobre o papel do jornalista, é importante lembrar da entrada do ciberespaço como ferramenta nas redações, pois o acesso da Internet evidenciou a necessidade deste profissional em considerar uma estrutura expandida, hipertextual, em rede, bem como de se pensar nos fluxos comunicacionais (CASTELLS, 1999). Com isso, fica mais evidente discorrer sobre alguns dos conceitos que entendemos como mais relevantes de Jornalismo digital, trazidos por Schwingel (2012):

² Sobrecarga de informação (tradução livre).

Tabela 1. Conceituando o Jornalismo Digital

AUTOR	CONCEITO
Elias Machado (2000, p. 19)	Jornalismo digital é todo produto discursivo que constrói a realidade por meio da singularidade dos eventos e que tem como suporte de circulação as redes telemáticas ou qualquer outro tipo de tecnologia por onde se transmitam sinais numéricos e que incorpore a intenção com os usuários ao longo do processo produtivo;
Alves (2004, p. 6)	Define o Jornalismo digital como todo processo discursivo que permite a multissequencialidade; que constrói a realidade por meio da singularidade dos eventos que podem ou não ser instantâneos e atualizáveis; que tem como suporte de circulação as redes telemáticas de alcance mundial ou qualquer outro tipo de tecnologia por onde se transmitam sinais numéricos; que incorpore a interação com os usuários ao longo do processo produtivo; que tenha a possibilidade de utilizar formatos em texto, áudio, imagem fotografada ou em movimento, sendo dessa forma, multimídia, o que demanda novos desenhos e retórica; e que, por fim, disponha de ferramentas que permitam a personalização do processo por parte de todos os atores envolvidos no processo de produção;
Salaverria (2005, p. 21)	Identifica o ciberjornalismo como a especialidade de Jornalismo que emprega o ciberespaço para investigar, produzir e, sobretudo, difundir conteúdos jornalísticos;

Assim, Schwingel chega em sua própria conclusão, relatando que

(...) ciberjornalismo é a modalidade jornalística no ciberespaço fundamentada pela utilização de sistemas automatizados de produção de conteúdos que possibilitam a composição de narrativas hipertextuais, multimídias e interativas. Os sistemas de gerenciamento e publicação de conteúdos são vinculados a bancos de dados relacionais e complexos (2012, p.21).

Com todos os conceitos apresentados acima, a definição que trazemos como norte para essa dissertação é a de que o Jornalismo digital, aqui entendido como

sinônimo de webjornalismo e ciberjornalismo, é a reconstrução da realidade vivida no cibermeio com a ajuda de composições como mídias, interatividade com o interagente, narrativas hipertextuais, personalização de conteúdos para cada público, utilização da memória e seus bancos de dados e, conseqüentemente, a atualização contínua.

Nesse caminho, conforme a maioria dos autores citados acima, o Jornalismo no ciberespaço possui como princípios básicos: “1) a multimídia; 2) a interatividade; 3) a hipertextualidade; 4) a customização dos conteúdos; 5) a memória; 6) a atualização contínua; 7) a flexibilização dos limites de tempo e espaço como fator de produção; 8) o uso de ferramentas automatizadas no processo de produção” (SCHWINGEL, 2012, p.21).

Em 2022 conseguimos perceber que há um crescimento exponencial quando falamos de comunicação digital interativa, principalmente porque é explorada toda a infraestrutura existente. Contudo, em 1995, de acordo com uma matéria da revista São Paulo (revista quinzenal que acompanhava o jornal diário impresso Folha de São Paulo), apenas 120 mil pessoas tinham acesso à Internet no Brasil (neste ano, o Brasil estava com 162 milhões de pessoas, ou seja, cerca de 7% da população total tinham a possibilidade de ter acesso à cibercultura³), mas na época, os principais jornais impressos, além das emissoras de televisão, Globo e SBT, já estavam na rede online com seus portais (ERCÍLIA, 1995). Em todo mundo e especificamente na América Latina, as empresas de comunicação com produtos impressos foram o setor da indústria que primeiro investiu no mercado Internet (BASTOS, 2002, *apud* SCHWINGEL, 2012). Hoje, em 2022, mais de 68 milhões de brasileiros possuem acesso a ela (com uma população de 215 milhões, segundo perspectiva da ONU, esse número representa cerca de 30% da população imersas nesta cibercultura, quase que 1% de aumento a cada ano⁴: não é à toa que cada vez aumenta a quantidade de conteúdos inseridos na Internet. Como trataremos nas referências de artigos, em média, 2,5 exabytes (1 seguido de 18 zeros) de dados são criados diariamente no mundo, e esse número dobra a cada 40 meses (MCAFEE, *et al.*, 2012, p. 1).

Segundo Palacios (2001), o jornalismo digital nasceu atrelado aos conceitos do impresso por conta dessa informatização das redações. Portanto, a historicidade do

³ IBGE.

⁴ Nota da autora.

Jornalismo digital o situa como uma evolução dos processos de digitalização das informações e de informatização das redações (VIANNA, 1992 *apud* PALACIOS, 2001).

Em 2001, Mielniczuk trouxe uma classificação histórica do jornalismo na web: 1º) a era da transposição de parte das matérias dos veículos impressos para a web (de 1995 a 1997); 2º) a era da metáfora do impresso (de 1997 a 1999), quando os produtos começam a apresentar serviços e informações com a utilização de características específicas da Internet, como links com chamadas para notícias de fatos que aconteceram no período entre as edições (utilização de hipertexto) e o e-mail e fórum de debates como interatividade com os leitores. Contudo, segundo Schwingel (2012), naquele momento, as entregas ainda não se distanciavam da estrutura e da representação do jornal ou da revista impressos. 3º) A era do Jornalismo de terceira geração, há 20 anos, quando os produtos e serviços são propostos de forma específica para a web (de 1999 a 2002), não só pelo avanço tecnológico dos microcomputadores, mas também pela possibilidade de transmissão mais rápida de sons e imagens (MIELNICZUK, 2001).

Barbosa traz uma quarta (2004) e quinta gerações (2013). A quarta contemplando a utilização dos bancos de dados inteligentes e dinâmicos para a elaboração de produtos do Jornalismo digital (de 2002 a 2013); e a quinta contemplando a expressão de convergência jornalística: quando, pelas tecnologias digitais implementadas de forma generalizada em diversos âmbitos, o profissional pode distribuir seus conteúdos de forma personalizada para cada público, por meio de múltiplas plataformas. Ao lembrar dos conceitos que norteiam o jornalismo digital, ambas já nos mostram a criação de algo único para o ciberespaço.

A quarta geração, por exemplo, traz luz à importância da memória no jornalismo. Quando acontece a busca no ciberespaço, essa procura ocorre dentro de um banco de dados, e a combinação dos resultados, ou do resgate da memória, acontece devido a algumas questões relativas à linguagem de comunicação computacional, juntamente com a humana, uma vez que o encontro das informações só ocorre quando não há ruídos nessa pesquisa.

Sobre esse processo entre o computador e o fator humano, Palamedi (2013, p. 78) lembra que McLuhan (1980) já alertava que “o meio é a mensagem”, quando exemplificou o uso da energia elétrica (e o resultado que é a luz) e o conteúdo da comunicação. A tecnologia serve como meio de comunicação entre usuário e

usuário, usuário e máquina, usuário e sistema ou mesmo usuário e projetista. O autor explica como funciona a linguagem computacional, que é diferente da humana, uma vez que ela é “baseada em comandos numéricos e instruções técnicas, normalmente de domínio exclusivo de engenheiros do produto, incompreensíveis a usuários comuns” (p. 65). Complementando com Thompson (1998):

Enquanto a interação face a face acontece em um contexto de copresença, os participantes de uma interação mediada podem estar em contextos espaciais ou temporais distintos. Os participantes não compartilham o mesmo referencial de espaço e de tempo e não pode presumir que os outros entenderão expressões denotativas (THOMPSON, 1998, p. 79 *apud* PALAMEDI, 2013).

Quando percebemos mais a fundo como ocorre a relação entre ser humano e máquina, conseguimos dividir em camadas: a mais baixa o maquinário ou o hardware, composta basicamente de equipamentos eletrônicos, entre eles unidades de armazenagem de dados, unidades de processamento de dados e unidades de comunicação de rede. A segunda camada é composta pelo software, responsável por orquestrar as atividades de todo o conjunto de hardware. A terceira camada é a interface, que torna sensível o ser humano à máquina, intermediando assim a interação entre máquina e o homem. (PALAMEDI, 2013).

Ou seja, presume-se que haja um diálogo direto entre a máquina e o operador da máquina (os usuários das plataformas), mas não há. Essa comunicação acontece entre aquele que deseja construir uma peça em um determinado formato e um projetista, que já estruturou o uso dessa peça para que seja usada no espaço-tempo para o projetista e o operador (PALAMEDI, 2013). Logo, o que tenta entender a comunicação humana é a interface. E se há algum ruído, ou seja, quando não é possível a consumação da mensagem, é necessário um maior envolvimento da comunicação com a usabilidade (PALAMEDI, 2013).

Apenas como uma amplitude da discussão e importância sobre o tema, na mesma linha, Haraway (2009) já criticava essa relação homem-máquina, quando trouxe à tona o questionamento de “poderíamos ser todos considerados ‘quimeras, híbridos –teóricos e fabricados - de máquina e organismo’; em suma, já todos somos ciborgues” (RITZMANN, 2012, *apud* HARAWAY, 2001: p.37). Sem esquecer que o manifesto de Haraway considerava muito mais as questões políticas que não caberiam nessa dissertação de mestrado, mas a autora levanta a ambiguidade do

relacionamento entre natural e artificial que as máquinas trouxeram no final do século XX. “Nossas máquinas são perturbadoramente vivas e nós mesmos assustadoramente inertes” (RITZMANN, 2012, *apud* HARAWAY, 2009: p.4).

A estruturação de um portal de notícias, por exemplo, exige a organização dos dados e exige um código visual, questões estruturais que não são do dia-a-dia do jornalista. Ferrari (2003) atentou para os passos fundamentais para o profissional que quer se aprofundar nessa parte, e uma delas foi a ferramenta de busca. Antes de 1996, era muito difícil encontrar qualquer informação na Internet se não dispusesse do endereço exato.

Com o advento dos sites de busca por palavras-chave, os usuários começaram a surfar pelas teias da grande rede e, com isso, cresceu a visitação aos sites. Endereços preciosos para os internautas, as ferramentas de busca ganharam um número gigantesco de usuários e deram origem ao formato portal (FERRARI, 2003, p. 30).

Como já tratamos, um dos principais desatadores de nós entre o Jornalismo impresso e o digital foi a implementação da Internet nas redações. Se em 2003 Ferrari vislumbrava que seria provável que os jornalistas comesçassem a “escrever notícias para vários formatos de distribuição, como Internet via cabo e móvel, televisão interativa e outros que irão surgir nos próximos anos” (FERRARI, 2003, p. 40), em 2022 esta situação já era realidade. A utilização de celulares, por exemplo, já é algo normalizado, não só para o profissional, mas também para todo ser humano que integra sua produção de conteúdo na *web*, que inclusive, faz parte da característica da cibercultura de ser interagente, avaliando, difundindo e compartilhando conteúdos (SANTOS, 2020, p. 6).

Um jornal virtual é a expressão máxima da realidade. Paradoxal, mas verdadeiro. Não se encerra, está sempre em movimento, é a cores, tem imagens, é global e instantâneo. “É a vida real. Não tem horas fixas, matérias predestinadas, páginas fechadas” (FERRARI, 2003, p. 46), esse conceito retratado em 2008 por Luís Delgado, diretor do *Diário Digital* continua sendo atual.

Mas Ferrari se atentou para outra questão também, pois além do jornalista ter que trabalhar com diversos tipos de mídias, é preciso desenvolver uma visão multidisciplinar, com noções comerciais e de marketing (2003, p. 40), que, de certa forma, historicamente, sempre existiu, mas agora é potencializada. E aqui pode ser o início da busca para se estar sempre na primeira página das plataformas de

pesquisa: na primeira página há mais acessos e, conseqüentemente, mais visibilidade para os canais de notícias.

Para Castells (1999), a sociedade não pode ser entendida ou representada sem suas ferramentas tecnológicas. Lèvy (1999) já havia constatado que as tecnologias intelectuais, suportadas pelo ciberespaço, amplificam, exteriorizam e modificam as funções cognitivas humanas. Dentre elas, os bancos de dados, hiperdocumentos e arquivos digitais de todos os tipos, a tecnologia favorece uma nova forma de acesso à informação (p. 158). Tecnologia e humanidade possuem entrelaçamentos cada vez mais comuns. Por exemplo, para encontrar uma informação, notícia, as pessoas recorrem mais às plataformas de busca (ou banco de dados), uma vez que as tecnologias já estão inseridas no cotidiano, seja ambiente de trabalho, educação, conhecimento, ou seja, há a “naturalização da tecnologia”. Ao adquirir informações em qualquer esfera (BARBOSA, 2008), o jornalista também se utiliza delas para uma base de dados. Esse profissional utiliza-as como um alicerce para suas atividades de pré-produção, produção, disponibilização/circulação, consumo e pós-produção (BARBOSA, 2008). Segundo Palacios (2011, p. 189), essas plataformas determinam o nível de facilidade oferecida pelo cibermeio para a busca do usuário por informações publicadas. “Compreende também a averiguação da existência ou inexistência de facilitadores de busca fora do cibermeio, como por exemplo, campo direto para o Google ou similares”. Em 2022, o Google (traremos um capítulo apenas sobre essa plataforma mais para frente) domina a preferência entre os usuários de ferramentas de buscas (NIKOLOVSKA, 2019). Em sua *homepage* na parte “Sobre”, explica que a missão dessa empresa é “organizar as informações do mundo para serem acessíveis e úteis para todos”.

Outra característica do cibermeio atrelado ao Jornalismo é a multimídia. Em 1999, Lèvy levantou a questão sobre o termo multimídia. Para o autor, ele é corretamente empregado quando são utilizadas diversas mídias para um mesmo tópico. Pelo exemplo do autor, é quando há o lançamento de um filme e, ao mesmo tempo, o lançamento de um videogame, exibição de uma série de televisão, camisetas e brinquedos.

Segundo Mielniczuk (2003) “no contexto do webjornalismo, a multimídia caracteriza a convergência dos formatos das mídias tradicionais (imagem, texto e som) na narração do fato jornalístico em um mesmo suporte” (p. 48). Em 2022 ambos os conceitos são utilizados. Um portal pode ter uma matéria sobre dado

assunto da forma escrita, com fotos ou vídeos, ou também, alocado na plataforma de vídeo, sendo divulgado nas redes sociais.

Já o termo hipertexto foi criado pelo filósofo e sociólogo americano, Ted Nelson⁵, (1963), para descrever como funcionava a plataforma de computação que ele utilizava para identificar partes de textos que levariam às fontes oficiais, chamada Xanadu. Tempos depois a expressão caracterizaria os links que conduziriam a outras páginas na Internet, como Ferrari (2003) sintetizou: “é um bloco de diferentes informações digitais interconectadas que, ao utilizar nós ou elos associativos (os chamados links), consegue moldar a rede hipertextual, permitindo que o leitor decida e avance sua leitura do modo que ele quiser, sem ser obrigada a seguir uma ordem linear” (p. 42).

Para as regras *SEO* (que veremos em um capítulo específico mais a frente), esses links denominam-se *backlinks*, e são de suma importância para o ranqueamento de um conteúdo na plataforma Google, pois eles traduzem a autoridade que uma dada matéria traz, devido aos links que possuem mais volume de acessos e estão atrelados a notícia primeira. Para Martins e Palacios (2016), os hiperlinks, ao apontarem para uma página na web, rede social ou outro aplicativo, conectam o usuário a um novo ambiente, que não necessariamente respeita o padrão de periodicidade da edição, por meio da mesma plataforma. Em 2011, Palacios já trazia a importância sobre tais links:

“São os links hipertextuais que estruturam, organizam e apresentam o grande volume de informações que pode passar a integrar a narrativa do fato jornalístico num cibermeio. Além disso, os links não são necessariamente criados a partir de relações estabelecidas um a um, como em sites estáticos.” (p. 48)

Carvalho, Lucas e Gonçalves (2010) trouxeram em um de seus artigos a questão sobre as mudanças fundamentais acontecidas após a proliferação exorbitante da web e os avanços na tecnologia moderna de computadores, constatadas por Baeza-Yates e Ribeiro Neto (1999, p. 7):

“1) Diminuiu-se visivelmente o custo de acesso a diversas fontes de informação; 2) Os avanços em comunicação digital permitem um melhor acesso às redes de trabalho locais e globais; e 3) A liberdade para incluir a informação que alguém considerou relevante contribuiu à popularidade da web”.

⁵ Para saber mais sobre Ted Nelson acesse <http://ted.hyperland.com>. Acessado em 29 de novembro de 2022.

A partir destas mudanças, os autores concluem que a web é altamente interativa, porque não é passiva perante tempo e espaço, e possibilita a criação de informações por diversas pessoas. A interatividade é uma das questões que aproxima o público do jornalista, pelo consumo do conteúdo por ele produzido. Incorporar essa comunicação no processo produtivo do profissional é algo que Alves (2004) já tinha inserido na definição do que é o ciberjornalismo.

Entre os recursos utilizados pelos jornalistas estão enquetes, fóruns, chats, e-mails, entre outros citados por Schwingel (2012). Ferrari (2003) tratou essa questão pelo ponto de vista da usabilidade, como um conjunto de características de um produto que definem seu grau de interação com o usuário (p. 60). Para ela, essa seria uma das principais iscas para prender a atenção do leitor na questão de conteúdo e na prática do e-commerce. Afinal, quem não se utiliza das avaliações de pessoas desconhecidas para definir a compra de algo pela Internet?

Outro modo de se utilizar dessa característica foi quando Ferrari (2003) fez diversos testes enquanto trabalhava em um portal de notícias para analisar se o volume de acessos aumentava. Entre algumas mudanças no design na *home*, a jornalista verificava por meio da interação quais chamavam mais atenção do interagente: “Ao promover a troca de chamadas e fotos nas primeiras páginas, percebi que os acessos cresciam e os leitores - monitorados diariamente pelas enquetes, fóruns e e-mails que chegavam às redações - estavam satisfeitos” (p.73).

Em 2022, todas essas características citadas do Jornalismo online são muito visíveis e, aliada a elas, estão ainda a atualização contínua e a flexibilização dos limites de tempo e espaço como fator de produção, uma vez que o profissional tem liberdade para a escrita da matéria de qualquer local, por diversos meios, como já havia previsto Ferrari (2003). Contudo, não só o jornalista se vale dessa flexibilidade, como também, qualquer interagente da web. Santos (2020) nos traz a questão de que os usuários dos veículos também possuem o poder de serem os emissores de conteúdos.

A informação começou a fluir de forma mais intensa dos usuários para os veículos através do recente poder dos primeiros em se tornarem, eles mesmos, emissores, capazes de produzir conteúdo em texto e audiovisual, utilizando na maioria das vezes apenas um celular e se tornando, em diversos arranjos, não apenas fontes, mas colaboradores, avaliadores, parceiros na difusão e compartilhamento do conteúdo e,

indiretamente, geradores de métricas e insights sobre preferências hoje indispensáveis em qualquer redação. (SANTOS, p. 6, 2020)

Contudo, os portais de notícias são os que possuem mais habilidade em gerar links das matérias. E o uso de ferramentas automatizadas no processo de produção possibilita essa criação em segundos. De acordo com Palacios (2011), os sistemas de gerenciamento de conteúdo (SGC) ou *Content Management Systems* (CMS), como o mais conhecido WordPress, passam a gerir também o processo de criação de *links*. “Quanto à aplicação da ferramenta, faz-se necessário esclarecer que a metodologia da ferramenta conjuga análise da home/portada da publicação e análise de matérias que correspondem à manchete principal na *home* para poder identificar as relações estabelecidas entre *links* propostos e os conteúdos a que remetem” (p. 463).

Juntamente com essa questão, ainda analisamos a parte de personalização de conteúdo. Não é de hoje que os jornais buscaram a transição do valor arrecadado do impresso para as assinaturas *online*. Segundo uma matéria feita pela Folha de São Paulo⁶ em 2020, foram 337.854 assinantes digitais pagos por mês, 3% a mais do que em 2019. O jornal O Globo ficou em segundo, com 332.176 (com 2% a mais que em 2019) e O Estado de São Paulo em terceiro, com 239.395 (com uma queda referente a 2019 de 1%). A assinatura possibilita que o jornalista faça personalização do seu conteúdo. Santos exemplifica que a conversão do produto físico para o digital é uma das ações desenvolvidas pelas empresas para se realinharem no cenário atual. Contudo, não é preciso pagar o acesso a um portal de notícia para perceber que a personalização acontece. Basta fazermos o *log-in* de uma conta Google de *e-mail*, continuar acessando *links* diversos para que na primeira busca na ferramenta de pesquisa comecem a aparecer produtos que você já havia procurado em algum momento. Ou então, conseguir personalizar o portal de notícias (sem pagar) para exibir na primeira página mais informações sobre o seu time do coração, porque foi questionado em algum momento por esse site qual era o seu time preferido de futebol masculino. Santos (2020) explica que essas questões são possíveis utilizando simples técnicas como o uso de palavras-chaves, *tags* (rótulos) e categorias.

⁶<https://www1.folha.uol.com.br/poder/2021/01/folha-encerra-a-decada-como-o-jornal-com-mais-assinantes-do-pais.shtml>. Acessa do em 30 de novembro de 2022.

Obviamente que tal tipo de serviço estaria disponível apenas para assinantes que depois de fazer seu *log-in* seriam identificados, possibilitando o carregamento do seu perfil de utilização e conseqüentemente a adaptação do conteúdo e de sua forma de apresentação aos seus interesses. Este é um bom exemplo do que chamamos de personalização estendida oportunizando a monetização, via venda de assinatura digital, potencializada por diferenciais de serviços e conteúdos exclusivos, incluindo material imersivo em realidade virtual ou agregado via realidade aumentada como descrito em Santos (2015). (SANTOS, p. 14, 2020)

Repensadas as principais características do cibermeio, atreladas ao Jornalismo, agora podemos seguir caminho para uma dentre todas essas questões que auxilia muito os profissionais da notícia, principalmente em se tratando da característica memória: a base de dados. Um dos principais portais que referencia essa organização das entregas de informação é o Google. Essa plataforma, em sua política de privacidade, reforça que a busca por ideias e produtos que frequentemente desafiam os limites da tecnologia existente.

2.1 A PLATAFORMA GOOGLE

O Google já foi referência em vários filmes, mostrando o dia a dia das pessoas que trabalham lá, geralmente, com uma cultura extrovertida e criativa, como em “Os estagiários” (2013) e “A ressaca” (2010). O nome vem de Googol, um número equivalente a um elevado a cem – ou seja, um seguido de cem zeros. É um número gigantesco, e isso representava, então, o número de pesquisas que o Google seria capaz de filtrar (WISE; MALSEED, 2007).

A empresa foi criada em 1998 pelos, então, estudantes de doutorado de Ciência da Computação da Universidade de Stanford, Larry Page e Sergey Brin, para a sua tese com o título *The Anatomy of a Large-Scale Hypertextual Web Search Engine*. Quase 25 anos depois, ela é uma das três mais valiosas empresas mundiais, segundo dados da revista *Exame* de fevereiro de 2022, valendo aproximadamente dois trilhões de dólares.

A criação da plataforma ocorreu por conta desses dois estudantes que queriam facilitar a sua própria busca em trabalhos acadêmicos. Na época, a Internet já estava bem difundida na Califórnia, mas ainda não contava com a organização de *Page Rank*. Tal tecnologia, criada por eles por meio de um algoritmo numérico, determina, até hoje, a relevância de um domínio, de acordo com a palavra-chave

pesquisada, entre outras regras, hoje conhecidas como *SEO*. Desde a sua solidificação, milhares de usuários consideram a Internet e o Google inseparáveis e indissociáveis, além dele ter mudado a forma de pesquisar, disseminar e recuperar informações, o que demonstra a capacidade dessa ferramenta (WISE; MALSEED, 2007).

Cabe lembrar aqui, que apesar de em 2022, quando se pesquisa pelo termo “*Page rank*” nas plataformas de busca, alguns sites, como Wikipedia⁷ (domínio público, com atualização humana de significados, comparado a um dicionário), descrevem como um algoritmo do Google, porém, tal técnica já havia sido utilizada em 1985 por Edmundo Landau, por conta da pontuação de jogos de xadrez. Na teoria dos grafos, a “centralidade do autovetor” é uma medida que influencia um nó em uma rede por meio de pontuações.

Cada nó contribui mais ou menos para conexão em questão. Ou seja, pontuações mais altas significam que são conexão mais coerentes. “Uma pontuação de autovetor alta significa que um nó está conectado a muitos nós que têm pontuações altas”⁸. Nessa comparação, o *Page Rank* utiliza a pontuação de referência de um site sobre dado assunto para a combinação de resultados na plataforma de busca.

Do lado de quem utiliza a plataforma, o Google criou política de privacidade e termos de uso. Aliás, como qualquer empresa que se utiliza dos dados de pessoas. Segundo o Serviço Federal de Processamento de Dados (Serpro), a política de privacidade de um site faz parte da estrutura de documentos para a proteção de dados. A política objetiva dar visibilidade ao tratamento de dados pessoais em um determinado serviço, atendendo princípios da Lei Geral de Proteção de Dados Pessoais (LGPD). É um documento endereçado aos usuários de um site, serviço ou sistema (titulares de dados). Ou seja, a intenção é explicar ao usuário daquele domínio, como os dados dele serão utilizados. O Google deixa bem clara essa questão, quando em sua página sobre a política utilizada, está escrito:

Quando você usa nossos serviços, está confiando a nós suas informações. Entendemos que isso é uma grande responsabilidade e trabalhamos duro para proteger essas informações e colocar você no controle. Esta Política de Privacidade destina-se a ajudar você a entender quais informações

⁷<https://pt.wikipedia.org/wiki/PageRank>. Acessado em 08 de dezembro de 2022.

⁸https://en.wikipedia.org/wiki/Edmund_Landau. Tradução livre. Acessado e 08 de dezembro de 2022.

coletamos, por que as coletamos e como você pode atualizar, gerenciar, exportar e excluir essas informações⁹.

O Google distribui em texto e vídeo as explicações, diminuindo com isso todos os ruídos possíveis sobre algum tipo de desentendimento, além de se utilizar de características do cibermeio em sua interface. Os tópicos trazidos pela empresa são:

1. INFORMAÇÕES COLETADAS PELO GOOGLE - Queremos que você entenda os tipos de informações que coletamos quando usa nossos serviços; Itens que você cria ou nos fornece; Informações que coletamos quando você usa nossos serviços; Sua atividade; Suas informações de localização.
2. POR QUE O GOOGLE COLETA DADOS - Usamos os dados para criar serviços melhores; Fornecer nossos serviços; Manter e melhorar nossos serviços; Desenvolver nossos serviços; Fornecer serviços personalizados, incluindo conteúdo e anúncios; Avaliar o desempenho; Entrar em contato com você; Proteger o Google, nossos usuários e o público.
3. SEUS CONTROLES DE PRIVACIDADE - Você tem escolhas em relação às informações que coletamos e como elas são usadas; Gerenciar, analisar e atualizar suas informações; Controle de privacidade; Modos de avaliar e atualizar suas informações; Exportar, remover e excluir informações.
4. COMPARTILHAMENTO DE INFORMAÇÕES - Quando você compartilha informações; Quando o Google compartilha informações.
5. MANTER SUAS INFORMAÇÕES SEGURAS - Incorporamos segurança nos nossos serviços para proteger suas informações.
6. EXPORTAR E EXCLUIR INFORMAÇÕES - Você pode exportar uma cópia das suas informações ou excluí-las da sua Conta do Google a qualquer momento.
7. RETENÇÃO DAS SUAS INFORMAÇÕES, COMPLIANCE E COOPERAÇÃO COM REGULAMENTADORES - Transferências de dados; Requisitos do Brasil.
8. SOBRE ESTA POLÍTICA - Quando esta política se aplica; Alterações nesta política.
9. PRÁTICAS DE PRIVACIDADE RELACIONADAS -Serviços específicos do Google; Outros Recursos úteis.

Já em seus termos de serviço¹⁰, o Google faz uma simples descrição sobre o que esperar da empresa, tais como seus produtos disponíveis:

Fornecemos uma grande variedade de serviços que estão sujeitos a estes termos, incluindo:

1. apps e sites, como a Pesquisa e o Maps;
2. plataformas, como o Google Shopping;
3. serviços integrados, como o Maps incorporado em apps ou sites de outras empresas;
4. dispositivos, como o Google Nest.

⁹ <https://policies.google.com/privacy?hl=pt-BR>

¹⁰ <https://policies.google.com/terms?hl=pt-BR>

Muitos desses serviços também incluem conteúdo que pode ser visto por streaming ou com o qual você pode interagir. Nossos serviços são feitos para operar em conjunto, facilitando a alternância entre atividades. Por exemplo, se um evento no Google Agenda inclui um endereço, você pode clicar no endereço para que o Maps mostre como chegar até ele.

Além, claro, da contrapartida do usuário:

Queremos manter um ambiente respeitoso para todos, o que significa que você precisa seguir estas regras básicas de conduta:

1. Obedeça às leis aplicáveis, incluindo controle de exportação, sanções e leis de tráfico de pessoas.
2. Respeite os direitos das outras pessoas, incluindo a privacidade e direitos de propriedade intelectual.
3. Não abuse nem prejudique terceiros ou a si mesmo, nem ameace ou incentive esse tipo de conduta. Por exemplo, enganar, fraudar, falsificar a identidade, difamar, fazer bullying, assediar ou perseguir terceiros.
4. Não abuse, prejudique, interrompa ou interfira nos serviços. Por exemplo, acessá-los ou usá-los de formas fraudulentas ou enganosas, introduzir malware, invadir, praticar spam ou ignorar nossas medidas de proteção ou nossos sistemas. Quando indexamos a Web para oferecer a você os resultados da pesquisa, respeitamos as restrições padrão de uso que os proprietários de sites especificam no código deles, então exigimos o mesmo quando terceiros utilizam nossos serviços.

Contudo, não é apenas em regras que a empresa trabalha nesse quesito. O Google também utiliza da característica do ciberespaço quando solicita que os usuários interajam, enviando denúncias quando verificarem algum rompimento dessas normas:

Se você encontrar outras pessoas que não estejam seguindo essas regras, muitos dos nossos serviços permitem que você denuncie casos de abuso. Se tomarmos providências com relação a uma denúncia de abuso, também forneceremos o processo descrito na seção: Tomando providências em caso de problemas.

Outro tópico abordado nos documentos do Google é a parte de “Tecnologia”. Aqui, a empresa deveria discorrer sobre como é feita a base tecnológica dela, como a parte de algoritmos, mas a empresa pouco fala sobre o assunto, deixando um único parágrafo discorrido nessa seção:

No Google, nós buscamos ideias e produtos que frequentemente desafiam os limites da tecnologia existente. No papel de empresa responsável, nós trabalhamos muito para ter certeza de que qualquer inovação esteja equilibrada com o nível adequado de privacidade e segurança para os

nossos usuários. Os nossos Princípios de Privacidade ajudam a orientar as decisões que tomamos em todos os níveis da empresa, de modo que possamos proteger e dar voz aos nossos usuários, enquanto cumprimos nossa missão contínua de organizar as informações do mundo.

Tais Princípios de Privacidade são:

1. Respeitar nossos usuários. Respeitar a privacidade.
2. Informar com clareza quais dados coletamos e por quê.
3. Nunca vender as informações pessoais de nossos usuários.
4. Facilitar o controle de privacidade.
5. Dar aos usuários a oportunidade de rever, mover ou excluir os próprios dados.
6. Integrar as melhores tecnologias de segurança aos nossos produtos.
7. Liderar pelo exemplo para melhorar a segurança online para todos.

Santos (2020) alertou que a utilização estratégica dos dados via algoritmos é o modelo de negócio do Google (e Facebook também). A decorrência disso, segundo o autor, é que gera uma percepção menor do valor do Jornalismo tradicional como produto, focado na personalização para fidelizar os usuários dos portais, por exemplo.

Hoje, a partir da lógica bastante explorada das bases de dados, é comum, entre os portais de notícia, gerar uma página personalizada apenas com o conteúdo de textos, fotos e vídeos sobre determinado time de futebol, por exemplo, utilizando técnicas simples como o uso de *tags* (rótulos), palavras-chave e categorias (SANTOS, 2020, p. 15). Inclusive, em nosso estudo de caso GE, veremos que, em uma das perguntas feitas, o coordenador explica que o acesso direto do *link* específico do time de futebol é muito maior do que a própria *homepage* do portal.

Ou seja, uma vez que essas plataformas (Google e Facebook) são gratuitas, e trabalham como grandes concentradoras de dados, as empresas de mídia foram obrigadas a portar seus conteúdos, produzidos com recursos próprios, para os canais digitais de difusão de informação controlados pelas empresas de tecnologia, como forma de minimizar a fragmentação das audiências que elas mesmas ajudaram a criar (SANTOS, 2020).

Como veremos no próximo subtópico, a varredura feita pelos algoritmos para extração dos dados para obtenção de resultados de uma certa pesquisa, tem como objetivo a classificação baseada para a coleta e organização desses dados, a técnica *PageRank* (MARTINS; PALACIOS, 2016, p. 156).

Essa técnica, segundo Holanda e Braz (2012), é uma fórmula matemática desenvolvida na Universidade de Stanford, nos Estados Unidos, que rastreia os sites em busca dos links gerados. O funcionamento ocorre da seguinte forma: o usuário realiza uma pesquisa na página principal do Google e as máquinas fazem uma busca no índice das páginas que correspondem e retornam os resultados que parecem ser os mais relevantes.

Tal relevância é julgada por mais de duzentos fatores, mas é o *PageRank* que administra toda a hierarquização dos resultados. Ou seja, para cada link gerado de uma página, em outro site é adicionado um *PageRank* ao site linkado, mas nem todos os *links* são iguais, o sistema identifica spams e outras ameaças ao resultado da pesquisa (GOOGLE, Central do Webmaster, 2010 citado por HOLANDA; BRAZ, 2012). Em linhas gerais, é importante lembrar que os resultados são listados, e para se ter um primeiro resultado, o *PageRank* serve para ordenar a importância que cada página tem na Internet.

Como os métodos utilizados pelo Google são identificados por meio dos termos de pesquisa dos usuários, quando uma página é analisada, verifica-se o que as outras páginas que estão vinculadas a esta tem a dizer sobre ela, preferindo as páginas nas quais os termos de busca estejam próximos entre si (HOLANDA; BRAZ, 2012; p. 11).

Uma questão importante também tratada por Holanda e Braz também, é que o Google ignora, nesse processo de busca, palavras e caracteres comuns, assim como ignora também dígitos ou letras isoladas, pois raramente ajudam na pesquisa e podem torná-la consideravelmente mais lenta. Segundo a empresa, refinar ou estreitar a pesquisa consiste no simples fato de se adicionar ou não palavras aos termos que já foram incorporados à busca. Ao adicionar mais palavras, seus resultados conterão um subconjunto específico das páginas retornadas pela sua consulta original (GOOGLE, Central de Ajuda, 2010, *apud* HOLANDA; BRAZ, 2012). Tudo isso melhora o resultado da pesquisa que em média, dependendo dos fatores de conexão, dura algo em torno de 5 segundos (MONTEIRO, 2008, p. 119 *apud* HOLANDA; BRAZ, 2012).

Conhecendo um pouco mais sobre a empresa Google, vale fazer uma linha do tempo sobre as principais mudanças da plataforma em seu algoritmo, que impactaram mais as regras *SEO*:

Imagem 2: Linha do tempo do algoritmo do Google



Gráfico desenvolvido pela autora, baseado no site <https://searchengineland.com/library/google/google-algorithm-updates>. Acessado em 30 de novembro de 2022.

Optamos por iniciar por 2011 devido a todas as pesquisas que foram feitas em cima de artigos entre 2012 e 2022, contudo, as atualizações do algoritmo do Google já datam desde 2003, com a primeira atualização denominada Flórida, implementada exatamente para iniciar com o *SEO* e acabar com sites de má qualidade que estavam sendo listados. Em 2011 a intenção foi de limitar a entrega dos sites que tinham baixa qualidade nos conteúdos e também com excesso de publicidade. Com a atualização Penguin, em 2012, a ferramenta penalizava quem se utilizava demais das palavras-chaves, na tentativa de manipular a entrega. Para isso foi utilizado o termo *keywordstuffing*. A partir de 2013, a atualização *Hummingbird*

colocou no foco da pesquisa a intenção do usuário, fazendo com que a pesquisa fosse mais precisa.

Em 2014 e 2015 passou a vigorar a obrigatoriedade do https e o fator *mobile* para que os sites comportassem a expansividade no tamanho da tela do celular, bem como a inteligência artificial entrou em ação. Já o algoritmo Fred fez uma atualização para que conteúdos mais bem desenvolvidos sejam entregues como prioridade. Em 2018, o algoritmo Medical veio para ajustar a qualidade de sites que impactavam na vida financeira ou de saúde das pessoas. No mesmo caminho em 2019, ocorrem diversas atualizações, mas as principais foram no propósito da confiabilidade e autoridade dos sites. Finalizamos em 2020, quando o Google retirou a duplicidade de entrega entre a posição “0” e a primeira posição nas pesquisas.

É importante frisarmos que as atualizações apenas se tornaram públicas após o algoritmo já ter aprendido e entrado em ação fazendo inclusive, com que diversos sites saíssem do ar ou perdessem posições na entrega.

2.2 SEO e TAG TITLE

A influência da tecnologia em diversos setores da sociedade é crescente, principalmente em se tratando de gerenciamento de dados. Como já vimos, indexação, armazenamento e recuperação já são ações normais quando a questão é a busca e inserção de informações na *web*. Desde o seu desenvolvimento, quando iniciada em um estudo por Tim Berners-Lee, no ano de 1989, a *web* tem sido aprimorada constantemente, tornando-se uma ferramenta popular considerada uma fonte valiosa de informação, disponibilizando textos, imagens, sons, fotos, vídeos, músicas, animações, multimídia etc (HOLANDA; BRAZ, 2012).

Já vimos que o Google, assim como outros motores de busca, foi criado com o intuito exatamente de facilitar essa pesquisa na *web*, além de organizar e tornar as informações acessíveis para todos, como estampa o *slogan* da empresa. Para Teixeira e Schiel (1997), a recuperação da informação tem três etapas: indexar, armazenar e recuperar:

Na recuperação de informação, a unidade básica de acesso à informação é a palavra-chave, pois, a partir desta unidade, diversos modelos podem ser desenvolvidos com o intuito de facilitar o acesso à informação e melhorar a precisão do resultado de uma consulta ou busca (BORSATO *et al*, 2006, p. 1).

A recuperação de informações na Internet (ou o resultado de uma busca por meio de uma palavra-chave) leva em consideração a URL, nome do site, metadados, como *TagTitle* e metadescrição, além de hiperlinks. Para Mercado (2002, p. 193), há alguns problemas a serem debatidos sobre a recuperação da informação na Internet:

- a) Informação e conhecimento utilizados como termos análogos – cada um tem o conhecimento adquirido de acordo com seu referencial pessoal. Ou seja, informação é uma questão, e ela tornar-se conhecimento, é outra etapa;
- b) Facilidade de dispersão – com diversos resultados da pesquisa, além de hiperlinks contidos nos textos, é comum que os usuários se percam na navegação;
- c) Perde-se muito tempo na rede – Dificuldade de se achar respostas corretas e confiáveis em pouco tempo, sem precisar passar horas abrindo todos os sites oferecidos pelo serviço de busca;
- d) Impaciência – acontece principalmente pela necessidade de mudar de um endereço a outro constantemente. Isso pode levar a aprofundar pouco as possibilidades em cada página encontrada. É difícil avaliar rapidamente o valor de cada página, porque há muita semelhança estética na sua apresentação;
- e) Conciliar os diferentes tempos dos usuários – alguns respondem imediatamente, outros, por fatores como habilidade ou familiaridade com as ferramentas, são mais lentos. Entretanto, a lentidão pode permitir um maior aprofundamento, e nesse sentido, na pesquisa individual esses ritmos diferentes podem ser respeitados.

O resultado na busca só acontece porque a página da *web* foi indexada, ou seja, a página foi assegurada de que seus documentos, seja texto, imagem, vídeo, etc, podem ser recuperados.

Com a indexação é possível se obter um aproveitamento melhor no processo de busca e recuperação da informação, uma vez que o elemento principal/fundamental estabelecido é a representação do conteúdo dos documentos (ARAÚJO JÚNIOR, 2007, apud HOLANDA; BRAZ, 2012).

Holanda e Braz (2012) descreveram como funciona a parte de indexação de uma página pelo Google quando ela é visitada pelo rastreador do Google (*Googlebot*). Segundo eles, em síntese, há uma análise de conteúdo e significado e, em seguida, armazenamento no índice do Google. As páginas indexadas podem ser exibidas nos resultados da Pesquisa Google se seguirem as diretrizes do Google, provenientes também, das atualizações dos algoritmos, conforme vimos no tópico anterior.

Além disso, Santos (2016, p. 150) lembra que ao visitar um site na Internet não estamos vendo a imagem de um único elemento completo, “mas sim o resultado da

construção feita pelo *browser* a partir de diversas partículas de informação”. Ou seja, vários arquivos enviados para o servidor em que o site específico está hospedado. Santos complementa explicando que depois disso, os arquivos em questão são “agrupados e estruturados pela ordem descrita no código da programação HTML”, que define, por sua vez, onde cada item que vemos deverá estar: texto, foto, título, vídeo e todas as multimídias.

Campos e Bax (2000) explicam que um dos principais fatores para se atentar ao ter um conteúdo postado na *web*, para que ela seja lida pelos Googlebots e indexada, são: estruturação do uso das palavras-chave, das *metatags* e de conhecimentos de HTML.

As *metatags* são linhas de código HTML, ou etiquetas, que, entre outras coisas, descrevem o conteúdo de um site para os mecanismos de busca. É nelas que se inserem as palavras-chaves. HTML (*HyperText Markup Language*) é uma linguagem de marcação utilizada para produzir páginas na Web, para serem interpretados por navegadores. É o código usado para transformar qualquer documento comum em documento acessível e disponibilizado à rede mundial de computadores e navegável por meio das ferramentas de busca de conteúdos na Internet. (HOLANDA; BRAZ, 2012).

É após esse rastreamento das páginas que é feita a formação de fila nos resultados. A organização é feita por meio do processamento das palavras encontradas e, segundo Holanda e Braz (2012), as informações que compõem o conteúdo são processadas e incluídas nos principais atributos e *tags* de conteúdo, como *TagsTitle*.

Com a tecnicidade que é utilizada para a inserção e busca de conteúdos na Internet, além do crescimento da massa de informação, torna-se fundamental a atividade dos jornalistas, que exercem o papel de filtragem e ordenamento. Mas também cabe a esse profissional que conheça as possibilidades das tecnologias em processo de aceleração e constante inovação de comunicação, fenômeno que está intrinsecamente vinculado à constituição do jornalismo contemporâneo (PALACIOS, 2003). Outra questão importante a se destacar, é que estar bem posicionado nos buscadores de extrema importância. Segundo Bertolini (2012), 50% das visitas aos portais de notícias vem por meio das plataformas como Google, os outros 50% são pelo tráfego direto, quando o interagente digita o endereço do portal direto no navegador.

Portanto, mesmo com tais tecnicidades nas práticas das regras *SEO*, devemos lembrar que elas auxiliam no processo do redator para que seu texto tenha maior visibilidade. Ou seja, como citam Pedrosa e de Moraes (2021), as técnicas de *SEO* podem ser entendidas como ações estratégicas do jornalista, para que sua notícia esteja destacada na *SERP* (página de resultados de busca).

Para isso, o profissional também deve entender que trabalha na porta de entrada do algoritmo do buscador, que tem por função, “rastrear e classificar diferentes sites em uma página de resultados em ordem decrescente de importância ou relevância” (PEDROSA, MORAIS, 2021, *apud* GILLESPIE, 2014).

2.3 MODELO DE ANÁLISE

Com base em todas as questões e autores que levantamos neste capítulo, fizemos uma tabela, de própria autoria, apresentando três das principais abordagens analisadas até aqui:

- As dimensões, que representam cada característica definida para uma análise, seja ela a Cibercultura, a Plataforma ou *SEO/ TagTitle*;
- Os indicadores que norteiam essa característica apresentada de forma tangível e demonstram o que significam cada estrutura;
- Os instrumentos, que são como essas características se apresentam para o interagente:

Tabela 2: Análise dos conceitos e suas principais abordagens

CONCEITO	DIMENSÕES	INDICADORES	INSTRUMENTOS
CIBERJORNALISMO	Hipertexto	Associação a outros textos já publicados no cibermeio	Links
	Interatividade	Série de processos diferenciados que ocorrem em relação à máquina, à publicação e às pessoas por meio do computador conectado à Internet.	Participação do Público, Enquetes, Concursos e Promoções, Fóruns, Notícias do Meio, Notícias, Chats, Vídeo Chats, Blogs, Consultórios, Fotografias, Vídeos, Comunidades, E-mail, Feed, Usabilidade, Acessibilidade.

	Multimídia	Utilização de várias formas visuais na mesma notícia	Fotografias, Vídeos, Áudios, Infográficos e agrupamento em galerias desses
	Design Homepage	Comprimento e esquema de estruturação da página inicial	Barra de rolagem, colunas, localização do menu/barra de navegação, número de itens informativos na página inicial, caixas ou frames, atributos de fundo, gráficos na página inicial, características da fotografia, aplicação de carrossel, elementos ou objetos multimídias, Integração de ferramentas gratuitas no desenho do cibermeio, Fios e linhas, Opções para criar diferentes homepages, Localização da publicidade, Tipologia da publicidade na homepage. Tipografia, considerações gerais, logomarca do veículo, títulos, antetítulos, bloco de textos, legenda da foto. Cores. Estrutura de Navegação e Interatividade
	Memória	Determina o nível de facilidade oferecida pelo cibermeio para a busca pelo usuário por informações publicadas.	Sistema de busca, localização de links
PLATAFORMA	Interface	Comunicação entre o sistema operacional e o usuário	Vídeos, YouTube, Meet, escrita, Drive, doc, áudio, hangout, grupos.
	Mercado	Empresa privada que oferece diversos serviços de forma gratuita	Gmail, GDrive, GAagenda, GDocumentos, Formulários, Hangouts e Grupos.

SEO	Algoritmos	<i>TagTitle</i> <i>Backlink</i> Palavra-chave Tempo de carregamento Indexação	Google Search Console Google Analytics
TAG TITLE	Marcação do código HTML	Caracteres do Pixels	Search Engine Results Page (página de resultados dos buscadores), Page Rank

Fonte: tabela elaborada pela autora (2022), baseado nos conteúdos dos autores Holanda, Braz, Santos e Schwingel.

Pela tabela podemos perceber que das dimensões apresentadas pelo ciberjornalismo, a memória e o hipertexto se interligam por meio de *links*. Uma vez que o resultado de uma pesquisa feita pela plataforma acontece, o indicador de memória determina a facilidade oferecida pelo meio para essa busca do usuário por informações. Seguindo essa linha, os algoritmos, que são os rastreadores, apresentam-se como dimensões do *SEO*, que possui como um de seus indicadores o *TagTitle*. E este último, por sua vez, apresenta-se como uma marcação do código de HTML, por meio do instrumento *SERP* e *PageRank*.

Tendo tais ligações em mente, passaremos para a parte de procedimentos metodológicos, os quais nos ofereceram mais informações para trilhar a importância desses conceitos para o Jornalismo.

3. PROCEDIMENTOS METODOLÓGICOS

Neste capítulo apresentaremos os procedimentos metodológicos utilizados. Anteriormente, definimos os principais conceitos abordados ao longo de toda a dissertação. Já neste ponto falaremos sobre o estado da arte, pesquisa feita por meio dos artigos achados; depois observamos um estudo de caso para que tais questões teóricas reunidas fossem melhores explanadas e delimitadas. Por fim, conversamos com um jornalista, por meio de uma entrevista, de um dos principais portais do Brasil, com filial no Paraná, o Globo Esporte (GE). Todos os procedimentos foram realizados com a intenção de se encontrar a resposta de como é feita a escolha das melhores estratégias pelo jornalista para que sua produção seja ranqueada na primeira página da ferramenta de pesquisa Google.

3.1 ESTADO DA ARTE

Nesta parte do processo, tem-se uma revisão do estado da arte, um tipo de revisão que tende a focar em assuntos contemporâneos e que pode oferecer novas perspectivas sobre um tema e apontar áreas que necessitam de mais pesquisa (GRANT; BOOTH; 2009, p. 95).

A revisão de estado da arte baseia-se na definição prévia de uma estratégia de busca, constituída de etapas. A primeira etapa da revisão consistiu em uma pesquisa em quatro plataformas distintas, sendo elas: Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI); Scopus e *Web of Science*. Nas plataformas foram realizadas buscas com as palavras-chave “SEO Jornalismo” (em português) e, para complementar a pesquisa, “*TagTitle*”, (nas línguas português, espanhol e inglês), ambas para artigos revisados por pares. Para todos os bancos de dados, os termos foram pesquisados nas seções de título, resumo e palavras-chave. Também houve um recorte temporal, de modo que foram incluídas apenas publicações entre 2012 e 2022.

Dentre os achados, tivemos apenas dois na palavra-chave “SEO Jornalismo”. Ambos na plataforma CAPES. O primeiro, de Jeferson Bertolini (2012), “Plataformas digitais: reconfigurações e heranças do título jornalístico”, traz uma comparação entre o jornalismo do meio impresso e o digital. Entre as técnicas, o autor relembra o

título, para o qual existia um limite de caracteres para que o diagramador conseguisse compor o texto dentro da página impressa. Já no jornalismo, ou na versão *online*, isso não ocorre. “Natural, então, que os títulos ganhem palavras a mais nas *homepages* e nas páginas internas de notícia” (2012, p.71).

No mesmo traçado sobre os títulos, o autor comenta que se o título ganha mais palavras, e é por meio dele que o interagente vai decidir se vai ou não ler o texto, então, “será que a tal gordurinha interfere neste processo?” Contudo, na época, o autor conferiu ao título um peso oito, numa escala de zero a dez, para os algoritmos do sistema. Em 2022 já sabemos, inclusive por meio dos artigos que traremos aqui sobre *TagTitle* para complementar esse levantamento, que essa *tag* acabou por não estar entre as principais ferramentas para o ranqueamento.

Bertolini confirmou, já em 2012, sobre a questão do processo de se acessar uma notícia no Brasil:

Do ponto de vista da audiência, estar bem posicionado nos buscadores é fundamental para qualquer site de notícias do Brasil. Isso porque, com alguma variação para mais ou para menos, 50% das visitas deles vêm por esses canais - o leitor prefere escrever palavras-chave em buscadores a digitar o endereço do site no navegador. Os outros 50% vêm de portais de referência (comum em grandes grupos de comunicação, quando um portal reúne seus jornais e rádios, por exemplo), tráfego direto (quando o internauta digita o endereço no navegador). (p. 72)

Outro dado trazido pelo artigo são as *pageviews* (visualizações de páginas da web) de alguns portais auditados pelo Instituto Verificador de Circulação (IVC). Em 2012, “com 21 sites conveniados, que juntos somam 700 milhões de *pageviews* por mês, mostra um avanço rápido do consumo da notícia via plataformas digitais e redes sociais no Brasil” (2012, p.72). Dez anos depois, em 2022, apenas o portal O Globo, um dos conveniados ao IVC, teve em outubro, mais 584 milhões (584,9 milhões) de visualizações, batendo, inclusive, um recorde.

A técnica do *lead* também é trazida à luz pelo autor, que relembra a importância da ligação do primeiro parágrafo (*lead*) com o título, citando Van Dijk (1990): ambos “compensam a grande desorganização que o leitor tem que desemaranhar ao longo do artigo” (p.73). Devido à grande quantidade de dados que o leitor possui em um texto na Internet, alguns mais outros menos importantes, segundo Bertolini, cabe ao leitor buscar a relação entre todas as informações contidas no texto (p.73).

Em suas considerações finais, Bertolini traz duas dicas para o jornalista manter a essência do título: 1) a de atrair o leitor para o texto; 2) a de dizer muito em poucas palavras. Além disso, o autor faz uma crítica entre essas duas questões dentro da web:

“Os jornalistas parecem não ter encontrado uma fórmula ideal para os títulos jornalísticos nas plataformas digitais. Falta encontrar o equilíbrio entre a característica histórica do dizer muito em poucas palavras e a neurose de escrever para os buscadores de notícias”. (BERTOLINI, 2012, p. 82)

Em nosso estudo de caso, veremos que, em 2022, o subtítulo além de servir como apoio para o que o título diz, é uma das *tags* que está metrificada para ranqueamento (conhecida pelos técnicos de *SEO* como metadescrição). E, na linha de raciocínio da citação de Bertolini acima, na entrevista feita para essa dissertação, também veremos que, nos tempos atuais, o jornalista dá mais prioridade para a escrita livre do que para as regras *SEO*, confrontando com a discussão já prevista pelo autor como uma questão a ser encontrada para o profissional.

No segundo artigo encontrado, de Pedrosa e Morais (2021), “Visibilidade Web nos Buscadores: Fatores Algorítmicos de *SEO On-page* (FAOPs) como Técnica e Prática Jornalística”, os autores trazem a importância do jornalista em tempos digitais, “devido a sua função social de produzir e compartilhar informações de interesse ao debate público” (p. 580), além de verificar que por meio do *SEO on-page*, foi possível a classificação dos Fatores Algorítmicos *On Page* (FAOPs), já que o fornecimento de dados de entrada (*inputs*) do primeiro é entregue aos algoritmos.

Pedrosa e Morais dividem as possibilidades do jornalista melhorar a entrega do seu texto em *On Page* e *Off Page*: fatores internos e externos que influenciam o ranqueamento na busca:

“Entre os fatores de *SEO* observados nas práticas de otimização de páginas, existem fatores de *SEO on-page* que podem ser controlados internamente pelo proprietário do site e de *SEO off-page* que dependem de variáveis externas (FISHKIN, 2019). Enquanto no *SEO on-page* realizam-se ações para otimizar o conteúdo de uma página web, no *SEO off-page* o foco está em estimular que sites externos cite os links desta página ou do domínio principal (LOPEZOSA *et al.*, 2019)”. (PEDROSA; MORAIS, 2021, p.581)

Já que os fatores *on-pages* são passíveis de mudança pelo próprio profissional, o artigo cita algumas das principais, mas também alerta que não há nenhuma certeza de que os textos aparecerão na primeira página de busca, contudo, o não cumprimento das regras é limitante para os conteúdos a serem ranqueados: retira um potencial material da *SERP* (*Search Engine Results Page*):

“Ser hiper-relevante a um tópico específico;
 Incluir o assunto abordado na *tag<title>*;
 Incluir o assunto abordado também na URL;
 Incluir o atributo de texto alternativo da imagem *<imgalt=“descrição da imagem”>*;
 Abordar reiteradamente o assunto no texto;
 Fornecer conteúdo exclusivo sobre o tópico;
 Ter link que retorne para seção daquele assunto;
 Ter link que retorne para a subcategoria daquele assunto (se aplicável);
 Ter link para a página inicial”. (PEDROSA; MORAIS, 2021, p. 582)

Com a descrição de um roteiro para os tópicos que o profissional consegue ajustar em seu texto, os autores complementam que, caso haja um nível de incompatibilidade entre a produção jornalística e as técnicas *SEO*, deve-se priorizar a escrita do Jornalismo (2021). Aqui temos a mesma questão trazida também por Bertolini (2012), e identificada em dois dos trabalhos que Pedrosa e Moraes ainda analisam do DigiDoc dentro desse artigo aqui citado: “*Posicionamiento Web y Medios de Comunicación: ciclo de vida de una campaña y factores SEO*” (CODINA *et al.*, 2017) e “*SEO y cybermedios: De la empresa a las aulas*” (LOPEZOSA *et al.*, 2019).

Ambos levantam uma clara diferença entre as rotinas de *SEO* desenvolvidas pelo redator e as do especialista em *SEO*. Para Pedrosa e Moraes, na visão dos profissionais de *SEO*, a função do redator é o conteúdo e a análise de tendências de busca, já a deles mesmos, seria assessorar os redatores com o intuito de otimizar o conteúdo para os buscadores.

Todos esses levantamentos feitos pelos autores serviram para que eles identificassem uma rotina melhor qualificada para o jornalista, a fim de que o passo a passo fizesse jus à escrita, mas também para a entrega do algoritmo. Eles denominaram de *Lide-FAOP*, remetendo a uma das principais técnicas do profissional, que responde às perguntas básicas em um processo de apuração e checagem de perguntas que deveriam constar na estrutura básica do primeiro

parágrafo de uma notícia, o *Lide*: Quem ou o Quê, Quando, Onde, Como, Por quê (2021). *Lide-FAOPs* identificados, conforme a lista a seguir (2021, p. 588):

1. As palavras-chave (Kw) aparecem nos demais fatores?
2. Os itens Multimídia (Mm) constroem a narrativa?
3. A URL (Ur) da página possui palavras-chave do tema?
4. O Título (Tt) foi elaborado para ser exibido no SERP?
5. Há metadados descritivos (Ds) sobre o teor da página?
6. A manchete jornalística ocupa o cabeçalho (H1)?
7. O texto âncora (Ac) leva a links com teor similar ou que complementam o conteúdo publicado?

Para complementar as perguntas deste lide algorítmico, propõe-se ainda uma sequência de questões de checagem antes da publicação, conforme proposto pelo *Framework SEO-RCP*:

- A. Há dupla titulação no material: título SEO <title> e título Jornalístico <h1>?
- B. A palavra-chave está distribuída ao longo do texto?
- C. Há conflito entre o manual/estilo jornalístico e alguma prática de SEO?
 - a. Se SIM, adota-se o estilo jornalístico e ajusta/retira o fator em conflito;
 - b. Se NÃO, o conteúdo está apto para publicação.

Além da contribuição com essa lista de checagem, os autores trazem luz ao que tentamos evidenciar com essa dissertação também. “Enquanto em outros países, já há um estudo avançado sobre tais técnicas, abordagens e checagens, no Brasil ainda falamos sobre a transição e adaptação do Jornalismo aos meios digitais” (2021, p. 589). Com isso, evidencia-se uma urgente qualificação na formação dos novos profissionais, com uma necessidade de uma disciplina em *SEO* para cursos de graduação, além da necessidade de fomentar mais estudos sobre a visibilidade da informação em meios algorítmicos (2021).

Já para a segunda base de dados, confrontando os resultados sobre, “*TagTitle*”, foram encontrados 160 trabalhos no total em todas as plataformas, excluindo duplicatas. Aqui, já foi possível constatar que o volume de produções envolvendo a regra de *SEO TagTitle* é pouco variado de acordo com o idioma das obras. Foi encontrada uma produção em espanhol, uma em português (pelo BRAPCI) e 158 em inglês.

Apesar do alto número de publicações em inglês já ser esperado, é importante também ressaltar que duas das bases de dados incluídas no escopo da pesquisa

(Periódicos da CAPES, BRAPCI e SciELO, esta última não teve nenhum registro de artigo com a palavra-chave, inclusive) são predominantemente utilizadas por autores brasileiros, o que supostamente colaboraria para um número maior de publicações em português. O fato de o número de publicações em inglês ser consideravelmente mais elevado pode apontar para um possível baixo volume de publicações brasileiras na interface explorada neste trabalho, de modo que ele também se propõe a ocupar esse espaço e fomentar mais discussões sobre o tema.

Os 160 artigos resultantes foram divididos em 35 categorias, sendo elas *Social Bookmarking* (12), Saúde (33), Bancário (01), Imagem (1), Otimização de Resultados (07), Imobiliário (1), Pesquisa (01), Turismo (03), Educação (1), Computação (1), Refatoração (1), Ruídos do Ambiente (1), Economia (2), LGBTQI+ (1), Vídeo (04), Organizacional (1), Design (4), Filosofia (5), Sociologia (7), Segurança (1), Sustentabilidade (1), Esporte (1), Multimídia (1), *Streaming* (1), Ecologia (1), Técnico Informática (5), Pedagogia (2), Publicidade (1), Indexação (1), Aplicativo (1), Fórum de Discussões (1), Mídia Social (37), Ranqueamento Híbrido (1), Geografia (16), Pesquisa (1), Política (1) e Categorização (1).

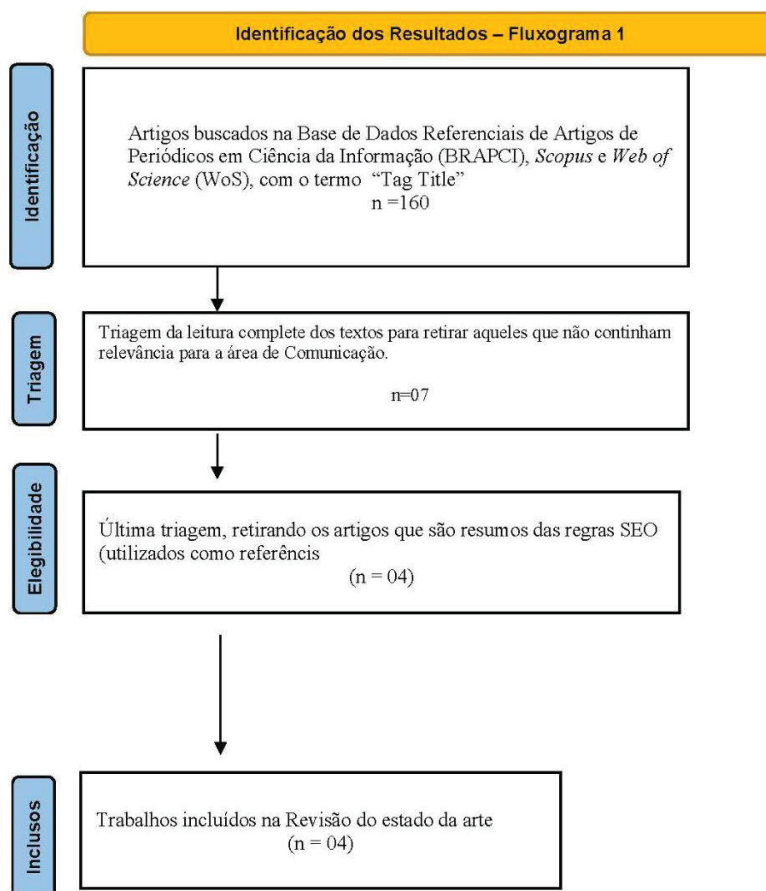
A segunda etapa da revisão compreendeu uma triagem por meio da análise dos títulos dos 160 trabalhos selecionados. Nesta etapa, foram excluídos trabalhos cujos títulos claramente apontavam não haver ligação direta com *TagTitle* como regra de *SEO*. A lista total de trabalhos encontrados e excluídos nesta etapa pode ser vista na Visão geral do Apêndice (p. 62 - os trabalhos marcados com “N” na coluna Relevância pelo título foram excluídos nesta etapa). Após as exclusões necessárias, restaram 25 trabalhos no total para análise textual.

A última triagem realizada nos trabalhos pesquisados consistiu na análise do texto completo dos últimos 25 trabalhos, o que resultou na exclusão de 18 deles pelo fato de as temáticas de *TagTitle* serem citadas fora do contexto a serem relevantes para o foco entre as regras e a utilização delas para a Comunicação. Por exemplo, foram retirados os artigos que tinham como estudo de caso ferramentas que se utilizam da abordagem híbrida, como a Social Bookmarking, CiteULike, mesmo sendo ferramentas que se utilizam de regras do *TagTitle* para entregarem os melhores ranqueamentos.

O motivo principal da exclusão é a de que elas são mais dependentes de uma interação humana do que de algoritmos, utilizando-se de folksonomia¹¹. Foram retirados também artigos que citavam a criação de um novo algoritmo para o ranqueamento ou para a utilização de alguma área específica como Turismo, Rede Social, Arquitetura, Imobiliário e Saúde.

Outro tipo de artigo que foi retirado para esse estudo foi o que apenas abordava um resumo das regras, contudo, desses (dos autores Pedrosa e Moraes, de 2021; KHAN; MAHMOOD, de 2018 e ZIAKIS, C. *et al* de 2019) utilizamos alguns conceitos e conclusões assertivas para este trabalho. Por tanto, retirando-se esses últimos três, tivemos quatro trabalhos restantes, que foram incluídos na Revisão do Estado da arte e serão discutidos a seguir. O Fluxograma 1 – Indicação de Resultados mostra um esquema simplificado do processo que levou à seleção desses artigos.

¹¹ N.A. É uma maneira que permite a cada usuário da informação a classificação com uma ou várias *tags*, em sua lista de “Favoritos”, para assim, compartilhar com quem tiver interesse a elas. [Folksonomy :: vanderwal.net](http://Folksonomy.vanderwal.net)



FONTE: a autora (2022), com base no modelo PRISMA utilizado para revisões sistemáticas.

Dentre os artigos que foram selecionados nesta dissertação, todos foram categorizados como sendo para Otimização de Resultados, sendo eles: An e Jung (2021), Song *et al* (2021), Siva Shankar *et al* (2020) e Lee *et al* (2016).

Dentre os achados, temos que em um dos primeiros artigos selecionados, de Song, Xiang; Zhu, Yi; Zeng, Xuemei; Chen, Xingshu, "*An Embedded-Based Weighted Feature Selection Algorithm for Classifying Web Document*" (2021) os autores relataram sobre o problema que gera o aumento da quantidade de textos dentro da Web para que eles sejam ranqueados de forma mais clara.

Um deles é que para tal processo, o código-fonte da Web enfrenta alguns problemas, pois ele demora a processar textos longos, outro é a vulnerabilidade a ataques cibernéticos e adulteração das páginas, e por fim, muitas distinções de categorias podem dificultar a seleção precisa de hierarquia na apuração da página.

Para a solução desses problemas, os autores propuseram-se a fazer um teste com metadescrições e *TagTitle* para diminuir esses chamados “ruídos” quando se entrega resultados na plataforma de pesquisa.

Isso só foi possível com a criação de um algoritmo, baseado em *Deterministic Finite Automata (DFA)*, que detectava e limpava os ruídos nas *metatags*, para que a classificação das escolhas dos textos fosse mais precisa. Segundo os autores, *TagTitle* e Descrição geralmente têm uma correlação mais forte com o tema da página da Web.

Outra constatação do artigo foi que o tamanho do texto da descrição é várias vezes maior do que o da *TagTitle*, logo, a descrição continha mais informações porque o texto geralmente é mais longo, o que ajuda o classificador a entender quando os dados não são suficientes. Como a *TagTitle* é mais curta, contém informações mais concisas, tornando mais difícil entregar dados suficientes que sejam bons para a classificação. Na conclusão do experimento, o algoritmo melhorou a limpeza da classificação hierárquica de dados, além de melhorar a entrega em razão da velocidade dos resultados.

Sabemos que o ranqueamento das páginas depende de vários fatores. Mas, o que já sabemos há mais de 30 anos, no Brasil sobre o jornalismo impresso e documentado por meio do “Manual de Redação da Folha de São Paulo”, em sua primeira edição, em 1984, a maioria dos leitores lê apenas o título dos textos editados. Por isso, ele é de alta importância. Ou o título é tudo que o leitor vai ler sobre o assunto ou é o fator que vai motivá-lo ou não a enfrentar o texto. Mesmo que, os autores ressaltem que a *TagTitle* contém pouca informação para que seja lida pelo algoritmo, a fim de que seja relevante na hierarquia, da parte do leitor, essa poderá ser a única coisa que ele poderá ler, mesmo nas primeiras posições da plataforma de busca. Aqui, já se entende que há uma importância para o algoritmo que não é a mesma dada a quem irá ler.

Aproveitando esse gancho, um dos artigos que não utilizamos para a contagem, de Ziakis, *et al*, “*Important Factors for Improving Google Search Rank*” (2019), trata de um Estudo de Caso e Revisão da Literatura. Pelos resultados da pesquisa do Estudo de Caso, os autores mostraram que os principais fatores que contribuem para classificações mais altas são a existência de certificado SSL do site, bem como a palavra-chave na URL, a quantidade de *backlinks* que apontam para outro um site relevante, o tamanho do texto e a idade do domínio.

Informações confrontadas com a Revisão de Literatura, que utiliza a *TagTitle*, metadescrição, entre outras regras como principais para as primeiras posições das pesquisas, e que, de acordo com os autores, não estão alinhadas. Utilizando essa linha de raciocínio, percebe-se que os algoritmos mudam muito rápido, conclusão tirada também pelos mesmos autores: “*Search engines nowadays change their algorithms regularly by adding and removing factors affecting the ranking of websites: SEO is a dynamic procedure*”¹² (2019, p. 2). Outro dado importante que traz neste artigo, são as taxas de cliques das primeiras posições: 30% (1ª posição), 16% (2ª posição) e 10% (3ª posição), enquanto a taxa de cliques para posições após a primeira página de SERPs (*Search Engine Result Page*) é limitada a menos de 2%.

No artigo de Shankar et al, “*An Embedded-Based Weighted Feature Selection Algorithm for Classifying Web Document*” (2020), propõem-se a criar um novo algoritmo, exatamente por conta do problema de muitos resultados em uma dada pesquisa, para otimizar o mecanismo de pesquisa para fornecer resultados rápidos e eficientes aos usuários. Aqui não foi utilizado o termo *TagTitle* como um dos principais, e sim, *headline* (mais conhecido como “H1”). A construção do novo algoritmo para ranquear de uma forma mais objetiva e assertiva é baseada em três etapas: *Wrapper*, *Filter* e *Embedded*. A primeira agrupa em subconjuntos todos os resultados, como se estivesse procurando os melhores agrupamentos dentre todos. Já a segunda, mede o grau que tem relação entre duas variáveis, detendo-se conceitos como o coeficiente de correlação de Pearson e análise de variância. E por fim, o último é um incorporado dos dois métodos. Um dos exemplos que os autores trazem para a eficiência do algoritmo por eles criado foi o de pesquisar a relação entre os links que cada texto da Web traz. Para ficar mais fácil o entendimento, colocamos a figura do próprio artigo para exemplificar:

¹²“Hoje em dia, os mecanismos de pesquisa mudam seus algoritmos regularmente adicionando e removendo fatores que afetam a classificação dos *sites*: SEO é um procedimento dinâmico”. Tradução livre.

Article Talk Read Edit View history Search Wikipedia

Vitamin A

From Wikipedia, the free encyclopedia

This article is about the family of vitamins. For the form usually used as a supplement, see Retinol.

Vitamin A is a **group** of **unsaturated nutritional organic compounds** that includes retinol, retinal, retinoic acid, and several provitamin A carotenoids (most notably beta-carotene).^{[1][2]} Vitamin A has **multiple functions**: **it is important** for growth and **development** for the maintenance of the **immune system**, and for good vision.^{[3][4]} Vitamin A is needed by the **retina** of the eye in the form of retinal, which combines with protein opsin to form rhodopsin, the **light-absorbing molecule**^[5] necessary for both low-light (scotopic vision) and color vision.^[6] Vitamin A also functions in a very different role as retinoic acid (an **irreversibly oxidized** form of retinol), which is an **important** hormone-like **growth factor** for epithelial and other cells.^{[4][7]}

In foods of animal origin, the **major** form of vitamin A is an ester, primarily **retinyl palmitate**, which is converted to retinol (chemically an alcohol) in the **small intestine**. The retinol form functions as a storage form of the vitamin, and can be converted to and from its visually active aldehyde form, retinal.

Retinol-Ester (Vitamin-A-Ester)
Retinol (Vitamin-A1)
Retinal (Vitamin-A-Aldehyd)
Retinsäure (Vitamin-A-Säure)

Figura 8 do artigo “An Embedded-Based Weighted Feature Selection Algorithm for Classifying Web Document”, p. 9. Demonstrando a importância do link (texto circular) acima do texto normal (quadrado), na pesquisa pela palavra *Vitamin A*, sendo este o resultado nº 1.

Pela imagem, percebe-se que o texto circulado (o qual contém um *link* para outro texto, um *backlink*), refere-se muito mais ao que estava sendo pesquisado (neste exemplo dos autores Vitamina A), do que o próprio texto normal, selecionado como quadrado. Em suma, os *backlinks* são mais relevantes do que palavras-chaves utilizadas ao longo do texto.

Pelas verificações dos autores, os resultados de todas as pesquisas que continham *links*, esses eram mais relevantes para um resultado mais limpo e objetivo, sem ruídos. Inclusive, os *links* contidos dentro dos resultados eram tão relevantes quanto o próprio resultado. O algoritmo de seleção criado com os recursos propostos utiliza-se dessas informações para a classificação.

Ou seja, mais um forte indício de que uma das principais questões a ser utilizada para estar nas primeiras posições de busca é a quantidade de *links* que o texto traz, deixando de lado, novamente, a questão do título, que é o primeiro contato do interagente.

Com mais essa informação, trazemos novamente o artigo de Ziakis *et al* (2019), quando os autores citam Zhang *et al* (2011), Chen *et al* (2011) e Patil Swati *et al* (2013). O primeiro realizou uma pesquisa comparativa sobre os fatores de *SEO* usando 116 sites, o segundo utilizou a regra *SEO* para mostrar como aumentar o tráfego de um site pode proporcionar um melhor relacionamento com o visitante do site e, o último que estudou como o algoritmo utiliza-se das regras *SEO* para classificar os resultados. Os três autores categorizaram em “*whitehat*” (chapéu

branco, tido como boas práticas) e “*blackhat*” (chapéu preto, práticas que rebaixam a classificação). Eles concluíram que, independentemente do país em que um site atende, os mesmos princípios de *SEO* são aplicados.

No mesmo caminho, Dean (2016), também citado no artigo de Ziakis *et al* (2019), fez uma pesquisa de *SEO* com o uso de um rastreador utilizando um milhão de sites. O objetivo principal da pesquisa foi comparar o ranking dos sites nas SERPs e as técnicas de *SEO* utilizadas neles. Os quatro estudos concluíram que os *backlinks* afetam principalmente a posição de um site seguido de conteúdo relevante, a existência de pelo menos uma imagem, bem como a existência de certificado SSL. No entanto, eles descobriram que técnicas como a existência da palavra-chave preferida na *TagTitle* não afetam o *SEO*.

No artigo de An; Jung, “*A heuristic approach on metadata recommendation for search engine optimization*” (2021), os autores trazem a importância de se estar nas primeiras posições da *SERP*, focados para os profissionais de marketing, que podem se utilizar das regras *SEO* para que sites sejam cada vez mais relevantes e tenham mais tráfego, além de ser uma importante ferramenta nos negócios.

Os autores trazem um dado que mais de 70% dos profissionais de marketing falham na realização de marketing online (*HubSpot*, p. 1). O estudo que eles se propõem é otimizar a metadescrição e as palavras-chaves na própria página do site de duas formas: em um primeiro momento apenas a otimização de metadescrições, num segundo a utilização de palavras-chaves, em grupos separados de sites e pesquisas.

Constatou-se que houve sim um aumento de tráfego quando combinados os dois juntos: palavras-chaves e metadescrição. A importância de trazer esse artigo para a elegibilidade da dissertação é a de que os profissionais de comunicação de marketing também podem se utilizar das regras *SEO*.

Para dar um pouco mais de luz a essa questão de otimização, trazemos um dos trabalhos que não contamos para a seleção de artigos, de Khan; Mahmood (2018). Os autores trazem a questão que com as palavras-chaves, tanto no conteúdo relevante quanto no título, juntamente com uma quantidade razoável de links de sites, ter-se-á uma melhor classificação nos resultados de pesquisa.

A disciplina de otimização de mecanismos de busca começou em meados da década de 1990 (KHAN; MAHMOOD, 2018, p. 1). Inicialmente, as táticas de *SEO* se restringiam a marcar as palavras-chave e as metadescrições, como trouxeram os

autores An e Jung (2021). A área de *SEO* começou a amadurecer a partir de 2003, quando foram introduzidas as estratégias de ancoragem de texto e interligação de páginas de sites. Mais tarde, em 2011, a inclusão de pesquisa vertical foi introduzida (ranqueamento).

E por fim, trazemos o artigo de Lee *et al*, “*Search engine optimization*” (2016), o qual fez-se um estudo aplicando as regras *SEO* às bibliografias da *Science Land* hospedadas pela *LG Sangnam Library*. Os novos conteúdos compatíveis com *SEO* foram aplicados aos motores de busca, procurando aumentar o número de visitas e identificar as principais técnicas de *SEO* que apresentaram os resultados mais promissores.

Os autores trazem um dado que exemplifica a grande importância de se ter cada vez mais um resultado que seja condizente à pesquisa: em média, 2,5 *exabytes* (1 seguido de 18 zeros) de dados são criados diariamente, e esse número dobra a cada 40 meses (p. 1). Agora, a quantidade de dados inseridos na Internet por segundo supera a adição anual de dados há 20 anos; só nos Estados Unidos há mais de 18 bilhões de consultas por mês (MCAFEE, *et al.*, 2012, p. 1). Com tanta informação, é necessário tempo para se filtrar e entender qual o melhor resultado para cada pesquisa. “*And time is a precious commodity in these-called “attention economy” of the Web where organizations compete with others to grab scarce commodity of users attention*”¹³ (HALAVAIS, 2013, p. 01).

Na investigação do artigo em questão, foram feitas algumas otimizações. Uma delas na *TagTitle* e outra na metadescrição, em ambas foram adicionadas palavras-chaves que tinham sentido com as bibliografias, além de se utilizar metadescrições para cada conteúdo distinto, antes o *ScienceLab* utilizava da mesma para todo site.

Como resultado comparativo, antes da aplicação de *SEO* em abril de 2014, a *Science Land* tinha 61.300 páginas indexadas no Google, 2.872 páginas no *Naver* e 45.500 páginas no *Daum*. Após, o *SEO* em setembro de 2014, os números aumentaram para 221.000 páginas no Google, 23.000 páginas no *Naver* e 53.600 páginas no *Daum*. Tal indexação levou, segundo os autores, ao aumento do tráfego pelos mecanismos de pesquisa e classificações mais altas de palavras-chaves.

¹³ “E o tempo é um bem precioso na chamada “economia da atenção” da Web, onde as organizações competem entre si para obter o bem escasso da atenção dos usuários”. Tradução livre.

Como conclusão, os autores criticam que houve uma limitação do estudo por conta da incapacidade de compreender o funcionamento interno dos algoritmos de classificação dos motores de busca, uma vez que são privados.

No estudo acima, comprovou-se a teoria de Ziakis *et al* (2019), quando comentam que para procurar qual o melhor resultado no ranqueamento, o rastreador procura informações sobre a estrutura de conteúdo de cada site e a melhor maneira de extrair essa informação é pesquisar principalmente pelas *tags*: metadescrição e *TagTitle*.

3.2 PESQUISA DOCUMENTAL GOOGLE

Appolinário (2009), no Dicionário de Metodologia Científica, descreve que a pesquisa documental se restringe à análise de documentos (p.152). A análise documental favorece a observação do processo de maturação ou de evolução de indivíduos, grupos, conceitos, conhecimentos, comportamentos, mentalidades, práticas, entre outros (CELLARD, 2008, p. 295). Com tais conceitos, intencionamos verificar os documentos sobre Política de Privacidade e os Termos de Usabilidade da plataforma Google. Imaginávamos que, por meio deles, poderíamos entender um pouco mais sobre os requisitos utilizados pela empresa na busca de resultados.

Os documentos utilizados foram encontrados por meio da própria plataforma de busca, Google, em seu buscador principal. As leituras se deram em períodos espaçados nos meses de maio e junho de 2022. O termo utilizado foi “política de privacidade Google” e no mesmo link como resultado “Google, Privacidade e Termos”: <https://policies.google.com>, foram encontrados os demais documentos utilizados nessa dissertação.

3.3 ESTUDO DE CASO

No estudo de caso, a metodologia utilizada foi a traçada por Yin (2010). Primeiro, é importante compreendermos que a intenção de se fazer esse estudo de caso é por conta do que trouxemos aqui, verificar como o jornalista utiliza as estratégias de SEO para ranqueamento por meio do estudo de um dos portais que mais entregaram como resultado na plataforma Google. Podemos entender também, que além dessa questão ser importante para os profissionais da Comunicação, é

também uma questão para qualquer pessoa ou empresa que anseia ficar na primeira página do Google, um fenômeno social complexo (2010). Outro fator que nos motivou foi o da revisão da literatura nos mostrar a falta de mais conteúdos sobre a questão, ela também nos ajudou a traçar um caminho para a análise do caso específico (2010).

Assim como Yin já havia comentado em seu livro, nossa intenção não será generalizar a partir da análise um caso. “Os estudos de caso são generalizáveis às proporções teóricas e não às populações ou aos universos” (2010, p.36), ou seja, não estamos concluindo que todos os portais são feitos da mesma forma, nem que deveriam seguir esse formato. Ao fazer esse estudo, nosso intuito foi expandir e generalizar teorias, como as que trouxemos nas análises anteriores, e não enumerar frequências (2010).

A primeira questão que tomamos como referência, como nos avisa Yin (2010), é classificar o tipo de questão da pesquisa a ser feito. No nosso caso, tomamos a questão “como”, pois estaremos analisando o formato que é inserido dentro das *tags Title* e *metadescrição* no portal Globo Esporte (ge.globo.com). O estudo foi feito nos dias 2, 4, 6 e 7 de novembro de 2022, gerando 44 links para análise. Tais links, seguindo o raciocínio de toda a dissertação, foi de analisar apenas os que apareciam na *home* do site, focando principalmente nos primeiros dez.

A essência de nossa análise foi realmente de lançar luz ao que alguns artigos nos trouxeram, com foco principal na desvalorização da *TagTitle* e na valorização de outras regras, como a da *metadescrição*. Yin explica que o estudo de caso é uma observação empírica que “investiga um fenômeno contemporâneo em profundidade e em seu contexto de vida real, especialmente quando os limites entre o fenômeno e o contexto não são claramente videntes” (2010, p. 39).

Como já retratou Barbosa (2008, p. 4), “cada vez mais as pessoas pesquisam pelas plataformas de busca”. Esse fenômeno que acontece na vida real, não está claramente explicado como ocorre, aqui, especificamente, pelo Google, como já vimos também.

Traçando nossa trajetória para a análise desse estudo de caso, vamos às questões delimitadas, segundo Yin (2010, p. 49):

- 1) A questão de estudo: Como o jornalista do portal GE, mais ranqueado pela nossa pesquisa, trabalha as *Tags*, *Title* e *metadescrição*, em seu portal de notícias?

- 2) Proposição, se houver: Tais *tags* são automatizadas? E pelos jornalistas? Segundo Yin (p. 50) é uma questão que pode ser examinada dentro do escopo do estudo, que podemos decifrar analisando as questões, nesse caso, tomamos como partida que, caso o Título da matéria seja igual a *TagTitle*, e o subtítulo igual à metadescrição, provavelmente o jornalista utiliza uma ferramenta automatizada para a construção do seu texto.
- 3) A unidade de análise: *TagTitle* e metadescrição. Para Yin, essa unidade é uma tentativa de definição que será analisado (p. 53)
- 4) Critérios para a interpretar as constatações. Aqui, Yin descreve que é a fase que o projeto deve dizer o que deve ser feito após os dados serem coletados (p. 57). Deixaremos para o final desse subtópico.

Ao final da dissertação, em Apêndice, deixamos a tabela de conteúdo do caso único, GE, para verificação. Utilizamos como fonte de evidência a documentação das notícias e os registros nesse arquivo como observação direta (p. 128).

Para melhor entendimento, fizemos um esboço dos dados relevantes para a análise que tentamos evidenciar com o estudo:

Título da <i>home</i>	1. igual a <i>tagtitle</i>	a. não faz pesquisa de volume e automatiza
Título da <i>home</i>	2. diferente da <i>tagtitle</i>	b. faz pesquisa de volume?

Para interpretar as constatações desse estudo, Yin nos ajuda com algumas técnicas e estratégias de análise. As que utilizamos foram as estratégias de análise conhecidas como Explicação rival simples (p.145), no qual se tem os resultados observados como uma consequência da análise de coleta de dados do estudo, tentando comprovar de fato essas evidências e não tentando rejeitá-las. Yin lista alguns tipos de rivais, e as utilizadas nessa dissertação são duas rivais artesanais ou “de ofício” (p.146):

1. A hipótese nula, no qual a observação é apenas o resultado de circunstâncias do acaso. Utilizamos essa pela falta de informações assertivas sobre o como é escolhida uma matéria para estar entre as principais na *home* do GE.

2. Viés da pesquisadora, que segundo Yin (p. 146), tem o efeito experimentador, a reatividade na pesquisa de campo. Estratégia escolhida para que a pesquisadora

pudesse analisar, durante a colheita de informações, quantos e quais dias seriam ideais para a análise.

Além disso, como uma estratégia indutiva (p. 142), utilizamos uma pesquisa qualitativa com uma pesquisa quantitativa: nesta fase, a utilização da questão de quantidade (dos links observados e analisados, tanto para verificar que o escolhido seria o portal GE, quanto dentro do próprio portal, a quantidade de links verificados) é utilizada para se ter garantias de duas questões, conforme indica Yin: os dados podem cobrir o comportamento e os eventos que estamos tentando explicar. Nesse caso, a hipótese seria a de que o jornalista utiliza a automatização em seus textos, ao invés de ajustar conforme as regras *SEO* para um melhor ranqueamento em cima das palavras que mais tem pesquisa. E a outra questão é a de os dados podem estar relacionados a unidade integrada de análise mais ampla, nesse caso as *tags*.

Dentro de uma estratégia, Yin considera a utilização de ao menos uma técnica analítica. Dentre as expostas por ele, escolhemos a técnica analítica de combinação de padrões (p. 146), que compara um padrão visualizado no empirismo, na coleta dos dados do estudo de caso, com os previstos antes da coleta.

Todos os dados do estudo de caso estão disponíveis também no link https://docs.google.com/spreadsheets/d/1ak_zsAQ6O0eICq3KZC-eON9ZpHkXLPou30vJLzVo4A8/edit?usp=sharing

3.4 ENTREVISTA

A entrevista foi realizada com o intuito de apresentarmos a visão do jornalista sobre as diversas temáticas aqui tratadas: ciberespaço, Google, *SEO*, além de como é o funcionamento do próprio portal GE, estudo de caso. Segundo Gil (2010, p. 44), as pesquisas descritivas são utilizadas para encontrar e dimensionar certa comunidade. Fizemos a entrevista no dia 1º de dezembro de 2022, às 15 horas. O entrevistado foi Guilherme de Paula Baptista, jornalista de formação desde 2007, coordenador do portal de esportes, GE, no Grupo Paranaense de Comunicação (GRPCOM). Apesar de o estudo de caso ser do GE nacional, a escolha pela entrevista do GE Paraná foi devido a proximidade da pesquisadora. A entrevista ocorreu por 30 minutos, pois o entrevistado estava atribulado de tarefa devido a Copa do Mundo, que estava ocorrendo. E por conta do tempo, não fizemos outras entrevistas. A escolha das perguntas foi em decorrência de todos os tópicos aqui

apresentados. As perguntas e a descrição completa da entrevista estão em Apêndice (p. 62).

4. RESULTADOS E ANÁLISES

Ao se levar em conta os artigos registrados nessa dissertação, vimos que quando o texto falava sobre *SEO* no Jornalismo, há ainda uma tentativa de ajustar o que foi feito na era impressa com o digital. Principalmente quando Bertolini (2012) faz uma crítica à falta de uma fórmula mais assertiva na escrita do jornalista, focando na parte da escolha do título, já que o profissional, muitas vezes, pode ficar em dúvida entre o dizer muito em pouco espaço, ou acertar na regra para o ranqueamento nos buscadores. Pedrosa e Moraes, quando elencam em 2021 algumas influências, sem certezas, dos fatores que podem auxiliar o jornalista a entregar melhor seu texto, não levam em consideração a automatização das plataformas de textos, como o WordPress que já insere automaticamente, desde aquela época, a palavra-chave na URL do *link* da matéria. Contudo, nesse artigo, há um questionamento interessante, que, conforme vimos em nossa entrevista, há de se estar levando em consideração pelas redações: o especialista em *SEO*. Técnico que pode auxiliar, e muito, o profissional de comunicação em seus textos, vídeos etc.

Já nos artigos especificados de *TagTitle*, houve uma melhor qualificação dos questionamentos sobre as regras de ranqueamento, principalmente descartando a *TagTitle* como uma das principais a serem utilizadas. Talvez isso se dê pela língua que é tratada, e, nesse caso, temos um possível desajuste nos artigos de língua portuguesa. Entre todos os analisados dessa segunda pesquisa, verificamos a importância dos links engajados para a matéria, o certificado SSL do site, e a importância da metadescrição para os algoritmos. Já nos de nossa língua nativa, ainda estamos discutindo sobre a importância de ter um técnico em *SEO* na equipe de jornalistas, ou seja, há um atraso de discussão que deve ser ajustado, uma vez que a informação é mundial na web.

Um dos locais que poderiam nos auxiliar trazendo mais informações sobre quais regras seriam as mais qualificadas para ranqueamento, os documentos do próprio Google, deixou a desejar, o que acontece é que mesmo que os algoritmos sejam “procedimentos codificados que, com base em cálculos específicos, transformam dados em resultados desejados” (GILLESPIE, 2018, p. 97), e que eles possuam início, meio e fim (PUIG, 2020), a busca no Google *Search* gera um fluxo de informação na *SERP*, também por meio da combinação de diferentes fatores de

SEO, que associa dados distintos tais como as palavras-chave utilizadas, o idioma pesquisado, a localização do endereço IP, o tipo de plataforma/navegador utilizado (desktop ou mobile), a autoridade daquele site sobre o tema, tudo isso para gerar um resultado orgânico a partir da execução algorítmica (PEDROSA; MORAIS, 2021). Ou seja, são variáveis que não nos permite traçar um padrão, e por isso, nos impedem de entender um pouco mais sobre os requisitos utilizados pela empresa na busca de resultados.

Já sobre a observação dos sites para se estabelecer um estudo de caso, o que percebemos é que, que as cinco primeiras notícias ranqueadas mudavam diariamente, depois, elencava-se a sessão das questões sobre o que as pessoas também perguntam e depois seguia para alguns sites que ficaram sempre listados: Wikipédia, Estadão, R7, UOL, Folha. Após a primeira semana, o site do GE começou a ficar em primeiro lugar, sem uma notícia principal, e sim, o site como um todo. Apenas para registro, quanto à verificação da utilização com e sem aspas na palavra de pesquisa Copa do Mundo, a entrega de conteúdo foi diferenciada pela questão de quantidade de resultados, cerca de 2,5% do total de resultado sem aspas era entregue na pesquisa com aspas:

Dias verificados	Média de resultados Copa do Mundo	Média de resultados “Copa do Mundo”
14	220.000.000	54.000

Fonte: a autora, 2022.

Além disso, a entrega de principais notícias diminuiu quando inseridas as aspas: de três para cinco, quando sem aspas. Era como se o pote menor da pesquisa com aspas coubesse no pote maior na pesquisa sem aspas.

No estudo de caso GE, temos 43 links de matérias levantadas na *home* do portal. Desses, quando analisados sobre o título ser igual ao *slug* (continuação da estrutura do link na barra de pesquisa, por exemplo, o texto em itálico, <https://ge.globo.com/futebol/times/palmeiras/noticia/2022/11/02/palmeiras-fatura-cerca-de-r-45-milhoes-com-premiacao-por-titulo-e-tera-extra-de-patrocinadora.ghtml> 18 eram iguais e 25 diferentes entre si. Já numa comparação entre o subtítulo e a *TagDescription*, 37 eram iguais entre si e 6 diferentes. Comparando título e *TagTitle*,

tivemos 19 iguais e 24 diferentes. E o *slug* em comparação ao *TagTitle*, obtivemos 30 iguais e 13 diferentes entre si. Tendo a hipótese de que se o código, nesse caso, as *tags*, são diferentes do subtítulo e título, a parte escrita do jornalista, significa que o profissional ajustou o código para uma melhor tentativa de ranqueamento. No estudo de caso tem-se que quase metade das matérias foram alteradas quando levamos em consideração o título. Já na questão da metadescrição, com *TagDescription*, e a parte textual, o subtítulo, existiram poucas mudanças, ou seja, o jornalista utilizou-se da automação da plataforma.

Aqui já podemos estabelecer uma conexão quanto à crítica de poucos estudos feitos na língua portuguesa. Pois se já tivéssemos levado em consideração todos os resultados obtidos pelos artigos feitos fora do Brasil, essa forma do jornalista aperfeiçoar os textos já estaria sendo de forma diferente, ou seja, eles estariam modificando mais a parte de metadescrição do que a de *TagTitle*, e já poderíamos ter outras hipóteses para melhorar ainda mais a entrega no ranqueamento dos textos.

E para finalizar, na entrevista feita com o coordenador do portal GE, verificou-se que, sim, é utilizada uma plataforma de automatização feita especificamente para o grupo Globo, e de que o jornalista tem a possibilidade de verificação das questões de volume de pesquisa, em tempo real. Além de que, constatou-se que o portal específico leva em consideração mais a *TagTitle* para ranqueamento do que a metadescrição, comprovada inclusive, no estudo de caso, conforme já observamos. Outra questão que trouxe a entrevista, é que o portal está olhando para os profissionais que se especializam em *SEO*, mesmo que, nesse caso, essa não seja um cargo específico, e sim, a junção com a função de jornalista.

5. CONSIDERAÇÕES FINAIS

Para se encontrar a resposta sobre como é feita a escolha das melhores estratégias pelo jornalista para que sua produção seja ranqueada na primeira página da ferramenta de pesquisa Google, praticamente seguimos uma tríade: autores que já escreveram sobre; observação de um portal de notícias e uma entrevista com um jornalista que trabalha nesse portal.

Como pudemos observar, o funcionamento da plataforma Google vai muito além das questões de mestadescrição e *TagTitle*, como uma abordagem simplista sobre regras de *SEO* pode induzir. Mesmo que apenas construído a partir de seis trabalhos, o estudo nos permitiu perceber que outros fatores colaboram para melhorar o desempenho em um site, quando o foco é o ranqueamento, principalmente quando visualizamos um referencial teórico consistente e complementar entre si.

Com os achados percebemos que o título em produções jornalísticas não é o fator mais importante para o buscador Google, contudo, vimos ainda, pela entrevista e pelo estudo de caso, que esse recurso editorial ainda é uma das técnicas que são mais alteradas para uma tentativa de melhor entrega pelo jornalista. Outro ponto importante, é que comparando ambas as pesquisas de estado da arte, uma, por ser mais focada na língua portuguesa, por trazer o termo “Jornalismo”, nos mostra que o campo nacional para esse tipo de assunto ainda é escasso, e mais, trata muito amplamente sobre as regras *SEO*. Podemos fazer uma analogia de que quando estamos doentes, seria como sempre fôssemos ao clínico geral, nunca em algum especialista sobre o assunto que vai esmiuçar cada detalhe dos sintomas. É notável o avanço dos artigos internacionais nesse quesito.

Neste sentido, essa dissertação atua não só no âmbito acadêmico, colocando mais luz ao afirmar que é necessário produzir e difundir mais conhecimento sobre as regras *SEO*, *TagTitle* em língua brasileira, especialmente nas bases de dados elegidas para este estudo, onde há pouca visibilidade do tema, mas também nos campos de mercado de trabalho e na estrutura de formação do Jornalismo, além de fomentar uma cidadania digital, focando na possibilidade de melhorar a visão crítica sobre o funcionamento das tecnologias digitais.

Sabemos que a inteligência artificial, em 2023, veio para ficar e ajudar o profissional de comunicação, e, talvez, mude as questões de ranqueamento dos

conteúdos pesquisados (nenhuma novidade), porém, a importância de se documentar cada fase em que passou o Jornalismo é de extrema importância, mesmo que haja uma caixa preta em uma empresa como o Google, dificultando a assertividade das escolhas a serem entregues para os interagentes do ciberespaço.

Outra questão que devemos lembrar é que os alunos que estão na academia devem ser os primeiros a ter tal tipo de informação, como as registradas aqui em 2023, uma vez que serão eles quem estarão à frente, num futuro próximo, no mercado de trabalho e, ou, pesquisando sobre as melhores formas de se alavancar os textos jornalísticos em alguma plataforma de pesquisa. Por tanto, essa dissertação não se configura como um registro finalizado. Trata-se de uma etapa para a sistematização dos estudos da área do Jornalismo, tanto no âmbito acadêmico, quanto no profissional.

6. REFERÊNCIAS

AN, Sojung; JUNG, Jason J. A heuristic approach on meta data recommendation for search engine optimization. *Concurrency and Computation: Practice and Experience*, v. 33, n. 3, p. e5407, 2021. doi: 10.1002/cpe.5407.

A ressaca. *Hot Tub Time Machine*, filme de 2010. Dirigido por Steve Pink.

BARBOSA, Suzana. Identificando remediações e rupturas no uso de bancos de dados no Jornalismo digital. In: *Anais do II SBPJor*. Salvador-BA/Brasil, 2004b.

BARBOSA, Suzana. Modelo Jornalismo Digital em Base de Dados (JDBD) em Interação com a Convergência Jornalística. *Textual & Visual Media*. *Revista de La Sociedad Española de Periodística*, v. 1, p. 1-308, 2008.

BARBOSA, Suzana. Jornalismo convergente e continuum multimídia na quinta geração do Jornalismo nas redes digitais. *Notícias e Mobilidade: O Jornalismo na Era dos Dispositivos Móveis*. Covilhã, UBI, LabCom, Livros Labcom. Universidade Federal da Bahia, Brasil, 2013.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. Modeling. In: *Modern information retrieval*. New York: Addison Wesley, 1999. Cap. 2. Disponível em: <http://www2.dcc.ufmg.br/livros/irbook/chapters/chap2.html>> Acesso em: 31 de agosto de 2022.

Bertolini, Jeferson. Plataformas digitais: reconfigurações e heranças do título jornalístico. *Vozes e Diálogo*. V. 11, n 2. Jul/Dez/ 2012. Edição especial, Pesquisa em Santa Catarina. Disponível em <https://periodicos.univali.br>. Acesso em: 31 de outubro de 2022.

BORSATO, G. G. et al. Recuperação de informação em situações de urgência-emergência no atendimento pré-hospitalar. *Revista Eletrônica de Sistemas de Informação*, v. 5, n. 3, 2006. Disponível em: Acesso em: 13 jun. 2022.

BOUTET, Charles-Victor; QUONIAM, Luc; SMITH, William Samuel Ravatua. Towards active seo (searchengineoptimization) 2.0. JISTEM – Journal of Information Systems and Technology Management, v. 9, n. 3, 2012. doi: 10.4301/S1807-17752012000300001.

BRAZ, Márcia; HOLANDA, Cíntia. INDEXAÇÃO AUTOMÁTICA DE CONTEÚDOS NA WEB: análise de sites de museus AUTOMATIC INDEXING IN CONTENT ON THE WEB: analysis of sites of museums, 2012. Disponível em <https://brapci.inf.br/index.php/res/download/50950>. Acessado em 10 de junho de 2021.

CAMPOS, F; BAX, M. P. Como os mecanismos de busca da web indexam páginas HTML. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 2000, Rio Grande do Sul. Anais eletrônicos... Rio Grande do Sul: FEBAB, 2000. Disponível em: <<http://www.bax.com.br/research/publications/ComoMaquinasBuscaIndexamPaginasWeb.pdf>>. Acesso em: 17 jun. 2020.

CARVALHO, Lidiane; LUCAS, Elaine e GONÇALVES, Lucas. ORGANIZAÇÃO DA INFORMAÇÃO PARA RECUPERAÇÃO EM REDES DE PRODUÇÃO E COLABORAÇÃO NA WEB. Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis, v.15, n.1, p. 71-86 jan./jun., 2010.

CASTELLS, M. A sociedade em rede. São Paulo: Paz e Terra, 1999.

CAVAVILHAS, João. "Hipertextualidade: Novas arquiteturas noticiosas". En CANAVILHAS, João (Org). WebJornalismo: 7 características que marcam a diferença. Covilhã (Portugal): Livros LabCom Universidade da Beira Interior. Págs. 3-24. ISBN: 978-989-654-144-6. 2014.

Codina L, Gonzalo-Penela C, Pedraza-Jiménez R, Rovira C. Posicionamiento web y medios de comunicación ciclo de vida de una campaña y factores SEO. Barcelona : Universitat Pompeu Fabra. Departament de Comunicació, 2017. 41 p. (Serie Digital Digidoc. Proyecto Comunicación Interactiva). DOI: 10.31009/DigiDoc.2019.inf23

CORRÊA E; BERTOCCHI D. A cena cibercultural do Jornalismo contemporâneo web semântica algoritmos aplicativos e curadoria. Matrizes, vol. 5, núm. 2, enero-junio, 2012, pp. 123-144. Universidade de São Paulo. São Paulo, Brasil.

CHEN-YUAN, Chen et al. The exploration of Internet marketing strategy by search engine optimization: A critical review and comparison. African Journal of Business Management, v. 5, n. 12, p. 4644-4649, 2011. doi: 10.5897/AJBM10.1417.

DEAN, B. Backlinko. Ranking Factors Study: Methods & Results. 2016. Disponível em: https://backlinko.com/wp-content/uploads/2016/02/Search_Engine_Ranking_Factors_Study_Methods__Backlinko.pdf. Acesso em: 24 jan. 2018.

DOS SANTOS, M. C. Data-Driven Journalistic Operation: Reshaping the Idea of News Values with Algorithms, Artificial Intelligence and Increased Personalization. Brazilian journalism research, [S. l.], v. 16, n. 3, p. 458–475, 2020. DOI: [10.25200/BJR.v16n3.2021.1295](https://doi.org/10.25200/BJR.v16n3.2021.1295).

DOS SANTOS, M. C.. A OPERAÇÃO JORNALÍSTICA GUIADA POR DADOS: reconfigurando a percepção de valor da notícia com algoritmos, inteligência artificial e personalização estendida. Universidade Federal do Maranhão, São Luís – Maranhão – Brasil. 2020.

DOS SANTOS, M. C.. NARRATIVAS AUTOMATIZADAS E A GERAÇÃO DE TEXTOS JORNALÍSTICOS: a estrutura de organização do lead traduzida em código. BRAZILIAN JOURNALISM RESEARCH (ONLINE) , v. 1, p. 160-185, 2016.

ERCÍLIA, M. Rede no Brasil anda bem devagar. Revista da Folha, São Paulo. p. 10, 31 dez. 1995. FAUS BELAU, A. La CienciaPeriodística de Otto Groth. Pamplona: Universidad Navarra, 1966.

FERRARI, P. Jornalismo Digital.FERRARI, Pollyana. Jornalismo Digital - 4ª edição. Editora: Editora Contexto Edição: 1ª. Idioma: Português ISBN: 8572442421, São Paulo, 2003.

GILLESPIE, Tarleton. The relevance of algorithms. *Media Technologies: Essayson Communication, Materiality, and Society*. v. 6, n. 1, p. 95-121, jan./abr. 2018.

GOUVEIA, Marco. Melhores Plataformas de Social Bookmarking. <https://www.marcogouveia.pt/melhores-plataformas-social-media-bookmarking>
Acesso em 07 de ago. de 2022.

GRANT, Maria J.; BOOTH, Andrew. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, UK, v. 26, p. 91-108, 27 maio 2009. Disponível em: <https://doi.org/10.1111/j.1471-1842.2009.00848.x>. Acesso em: 14 jun. 2021.

HARAWAY, D. (2000): “Manifesto Ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX”, em T. T. Silva: *Antropologia do ciborgue: as vertigens do pós-humano*, Belo Horizonte, Autêntica.

KHAN, M. N. A.; MAHMOOD, A. A distinctive approach to obtain higher page rank through search engine optimization. *Sādhanā*, v. 43, n. 3, p. 1-12, 2018. doi: 10.1007/s12046-018-0812-3.

KOVACH, BILL. ROSENSTIEL, Tom (coautor). *Os elementos do Jornalismo: o que os jornalistas devem saber e o público exigir*. 2.ed. São Paulo: Geração Editorial, 2004.

LEE, Sungin et al. Search engine optimization: A case study using the bibliographies of LG Science Land in Korea. *Library Hi Tech*, 2016. 10.1108/LHT-02-2016-0014.

LEVY, P. . *CIBERCULTURA*. São Paulo: Ed. 34, 1999.

MAIMONE, Giovana; SILVEIRA, Naira. *Cognição humana e os Paradigmas da Ciência da Informação*. Pontifícia Universidade Católica de Campinas, São Paulo. *Revista Eletrônica Informação e Cognição*, v.6, n.1, p.55-67, 2007. ISSN:1807-8281.

MARTINS, Eduardo. Manual de Redação e Estilo do Estado de São Paulo. 3ª edição, revista e ampliada — São Paulo: O Estado de S. Paulo, 2001.

Mark A. Tietjen; Robert M. Myers. "Motivation and job satisfaction". Palm Beach Atlantic College, West Palm Beach, Florida, USA. Management Decision 36/4. 1998, 226–231. MCB University Press ISSN 0025-1747.

Mcafee, A.; Brynjolfsson, E. (2012). Big Data The Management Revolution – Exploiting vast new flows of information can radically improve your company's performance. But first you'll have to change your decision-making culture. Business Review, edição outubro.

MELO, José Marques de. Gêneros Jornalísticos. In: MELO, José Marques de; ASSIS, Francisco de. Gêneros jornalísticos no Brasil. São Bernardo do Campo: Universidade Metodista de São Paulo, 2010.

MERCADO, L. P. L. A Internet como ambiente de pesquisa na escola. In: MERCADO, L. P. L. (Org.). Novas tecnologias na educação: reflexões sobre a prática. Maceió: EDUFAL, 2002. p. 191-204.

NIELSEN, Jakob, R. Molich. 1990 "Heuristic valuation of user interfaces," Proc. Of CHI'90, ACM, New York.

NIKOLOVSKA, Hristina. 60+ SEO Statistics to Help You Rank #1 in 2021. SEO Tribunal, 2019. Disponível em: <https://seotribunal.com/blog/stats-to-understand-seo>. Acesso em: 6 ago. 2022.

OS estagiários, The internship. filme de 2012, dirigido por Shawn Levy.

OLIVEIRA, R. R. Entre híbridos e ciborgues: as ficções anti-modernas de Bruno Latour e Donna Haraway. VI Simpósio Nacional de História Cultural Escritas da História: Ver – Sentir – Narrar. 2012, UFPI. Anais... Teresina, 2012. In: p. 1-10.

PAGE, M. J. et al. The PRISMA 2020 statement: a updated guideline for reporting systematic reviews. Syst Rev, v. 10, n. 89, 2021. 10.1186/s13643-021-01626-4.

PALACIOS, Marcos. (1999) O que há de (realmente) novo no Jornalismo online? Conferência proferida por ocasião do concurso público para Professor Titular na FACOM/UFBA, 21.09.1999

PALACIOS, Marcos; MIELNICZUK, Luciana. Considerações para um estudo sobre o formato da notícia na Web: o link como elemento paratextual. Trabalho apresentado no XI Congresso da Compós. Brasília, 2001

PALACIOS, Marcos. Ferramentas para análise de qualidade no Ciberjornalismo. Volume 1. Covilhã, UBI, LabCom Books. 2011.

PALACIOS, Marcos. Ferramentas para análise de qualidade no Ciberjornalismo Volume 2. LabCom IFP. 2011.

PALACIOS, Marcos. Jornalismo Online, Informação e Memória: Apontamentos para debate. Informação e comunicação online. Jornalismo online, Covilhã, Universidade da Beira Interior, Portugal. 2003.

PALAMEDI, F. R. . "A usabilidade como instrumento da análise da função comunicativa e interfaces digitais", Júnior & Dos Santos (Orgs). Comunicação, Tecnologia e Inovação: Estudos Interdisciplinares de um campo em expansão. Porto Alegre (Brasil): Buqui. Págs. 63-85. 2013.

PATIL SWATI, P.; PAWAR, B. V.; PATIL AJAY, S. Search engine optimization: A study. Research Journal of Computer and Information Technology Sciences, v. 1, n. 1, p. 10-13, 2013. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1070.1729&rep=rep1&type=pdf>. Acesso em: 6 ago. 2022.

PEDROSA, L.; MORAIS, O. J.. Visibilidade Web nos Buscadores: Fatores Algorítmicos de SEO On-page (FAOPs) como Técnica e Prática Jornalística. Estudios sobre El Mensaje Periodístico, v. 27, n. 2, p. 579-591, 2021.doi: 10.5209/esmp.71291.

RITZMANN, R. . Entre Híbridos e Ciborgues: As ficções anti-modernas de Bruno Latour e Donna Haraway. In: VI Simpósio Nacional de História Cultural, 2012, Teresina. Anais do VI Simpósio Nacional de História Cultural Escritas da História: Ver - Sentir - Narrar. Uberlândia: GT Nacional de História Cultural, 2012.

SILVA, Carlos Eduardo Lins da. Manual de Redação da Folha de São Paulo. 2005. Disponível em: https://www1.folha.uol.com.br/folha/circulo/manual_edicao_t.htm. Acesso em: 7 ago. 2022.

SIVA SHANKAR, G. et al. Na embedded-based weighted features election algorithmfor classifying web document. Wireless Communications and Mobile Computing, v. 2020, 2020. doi: 10.1155/2020/8879054.

SCHWINGEL, C. Ciberjornalismo. São Paulo: Paulinas, 2012.

SONG, Xiang et al. Hierarchical contaminated web Page classification based on meta tag denoising disposal. Security and Communication Networks, v. 2021, 2021. doi: 10.1155/2021/2470897.

TEIXEIRA, C. M. S.; SCHIEL, U. A Internet e seu impacto nos processos de recuperação da informação. Ciência da Informação, v. 26, n. 1, 1997. DOI: 10.18225/ci.inf..v26i1.754 Acesso em: 17 de outubro 2022.

WISE, David A. e MALSEED, Mark. "Google a História do Negócio de Mídia e Tecnologia de Maior Sucesso dos Nossos Tempos. Edição Português. 1ª edição Rio de Janeiro: Rocco, 2007.

YIN, Robert K. Estudo de caso: planejamento e métodos. 4 ed. Porto Alegre: Bookman, 2010.

WANG, Fuxue; LI, Yi; ZHANG, Yiwen. Na empirical study on the search engine optimization technique and its out comes. In: International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2, 2011.

IEEE, 2011, p. 2767-2770. doi: 10.1109/AIMSEC.2011.6011361.

ZIAKIS, C. et al. Important Factors for Improving Google Search Rank. Future Internet, v. 11, n. 2, p. 32, 2019. doi: 10.3390/fi11020032

7. APÊNDICE

7.1 QUESTIONÁRIO

7.1 Perguntas da Entrevista

7.1.1 Identificação

a) Qual seu nome?

Guilherme de Paula Baptista

b) Qual sua formação profissional?

Sou formado em jornalista desde 2007

c) Em qual instituição você se formou?

Universidade Tuiuti do Paraná.

d) Em qual ano você se formou?

2007

e) Qual empresa você trabalha?

Coordenador de esportes na RPC

f) Há quanto tempo trabalha nesta empresa?

7 meses em dezembro

g) Sua empresa é nata digital ou é migrante do impresso para o digital?

TV para digital ge.com/parana é o nosso braço digital aqui na TV.

7.1.2. Cibermeio

a) Como é a rotina da equipe do portal do Globo Esporte para a produção de notícias? 2:10 - 4:49

Hoje a equipe é completamente integrada, trabalha com a mesma coordenação da TV e do site sob o mesmo guarda-chuva, e estamos num processo cada vez mais forte de integração dessas mídias. Hoje a gente trata do projeto Globo Esporte. Claro que tem afinidade de função, repórteres mais dedicados ao site. Mas a ideia é que todos participem de alguma forma de todo processo. Desde o site até a TV. Hoje é muito comum para nós o cross de conteúdo, hoje a gente trabalha com conteúdos que migram da TV e vão a uma versão estendida do site ou uma versão estendida do site que viram uma versão compacta para a TV. A produção ela acontece na mesma maneira, o nosso ideal, a gente está numa transição de gestão construindo isso, a gente está numa ideia de que todos produzam conteúdo para os dois produtos maiores, TV e site, de uma maneira organizada e coordenada, mas que todo mundo participe de alguma forma do que a gente tem de conteúdo no site e também na TV. Claro, temos alguns repórteres dedicados mais a TV, mas também tem essa multitarefa de entrega no site. Por exemplo, a Nadja Maua que é repórter

da TV que tem um blog, que é especializado em mercado do futebol, algo muito forte, que as pessoas procuram, e ela tem uma característica e habilidade de trazer essa notícia em primeira mão. É o símbolo dessa multitarefa, o repórter que caminha no nosso digital, no caso do site, nosso portal, e também é repórter da TV, então trabalha com conteúdos diferentes nas duas plataformas, digamos assim.

b) Mas é mais fácil uma pessoa vir da TV pro portal do que o contrário?

É, ainda é, porque a gente, pega uma matéria, vou dar um exemplo da Libertadores, a gente teve uma série que apresentava a família dos jogadores do Atlético que iam jogar a final contra o Flamengo, a gente pegava esse perfil que na TV era um VT de 5 min, e transformava num grande perfil, aí ampliavam com texto, um texto mais robusto, mas o conteúdo da TV também estava lá, então, ele poderia assistir e se ele quisesse se aprofundar sobre perfil daquele do jogador, ele poderia ler o conteúdo feito para ele. O que a gente tenta fugir é que uma coisa não se torne apenas uma repetição da outra, tem que ser versão, ou uma versão estendida, ou uma versão adaptada para a plataforma, mas tem que ser algo inédito, algo novo para ser consumido naquela mídia.

c) Sempre pensando que a pessoa vai consumir nos dois, isso?

Exatamente essa é a nossa ideia.

d) Como é feita a definição da notícia? 6:00

Isso é muito diferente da TV e portal, gosto de traçar um paralelo que a gente trabalha sempre com os dois produtos. Mas aqui vale traçar um paralelo pra explicar. Na TV, a gente tá na TV aberta, entre dois jornais que tratam de notícia gerais, o público é muito amplo, muito grande, então ali a gente precisa tratar mais de grandes histórias que consiga atrair a pessoa que assiste Meio Dia Paraná e continua no Globo Esporte, eu preciso achar uma forma para tratar com pessoas que também não gosta de esporte. É o nosso grande desafio. No site não, no nosso portal é um nicho, a pessoa vai lá procura o ge.globo.com/parana para encontrar as notícias do futebol paranaense, mais do que isso, isso é muito mais, muito mais mesmo. A pessoa procura só a home, só a página do clube que ela torce, então, se ela quiser só consumir notícias do Atlético no nosso portal, ela tem como ir na Internet, o conteúdo segmentado é uma tendência, não só no futebol, hoje cada vez mais as pessoas vivendo em bolhas, então a pessoa vai viver numa bolha só sabendo notícias do time dela. No orgânico mesmo, se ela procurar no google "GE Atlético", ela vai cair direto na página do Atlético, claro, no orgânico ela talvez possa cair no conteúdo recomendado, né, mas se ela quiser ela consome só conteúdo do Atlético, do Coritiba, principalmente que são os dois carros chefes do portal.

e) Reuniões de pautas não existem mais?

Hoje a gente tem muito mais um alinhamento do que a gente quer dar como conteúdo, mas dentro deste alinhamento. Esse alinhamento é que dá a autonomia para que os repórteres toquem a pauta. Tem direcionamento de setorista por semana, nessa coisa de mercado um pouco menos, mas geralmente um setorista

que cuida do Atlético, do Coritiba e um que cuida do interior. Mas nós estando alinhado com o tipo de conteúdo que a gente quer, a gente quer ter no portal grandes histórias, grandes personagens, que isso carrega o nosso conteúdo para a home nacional que traz um impacto significativo de audiência também. Ele sabendo, que precisa ser feito, estando alinhado, a missão é alinhar, aí a produção é uma equipe que trabalha de forma autônoma pela organização que ele tem. Mas a gente tem reunião de pauta semanal também, quando está com nossa rotina certinha, ou em grandes momentos (10:06).

f) Como é feita a decisão para o que vai para a home do portal?

O estado é a gente quem define, então nós temos aqui várias homes, a home do GE PR, que é a página /parana, então, essa é uma home do futebol paranaense. Ali a gente sempre deixa, além do que a gente considera mais quente, mais forte, o que vai trazer mais acesso, também sempre carrega algo do interior ali, porque as pessoas ainda não tem o costume de procurar a home do Londrina, do Operário, por exemplo, que são clubes que tem a home, mas que são homes menos procuradas, então, a gente deixa na home do GE PR, a dos clubes, é de acordo com o giro de notícias que a gente vai fazendo, então, meio em cima da atualização, claro, às vezes acaba de deixar alguma que foi super bem ali, porque se ela tem bom desempenho, não tem o porquê dela ser retirada. O desempenho é pelo número de acessos, a gente tem o controle pelo tempo real, como está o desempenho por notícia. Isso é muito claro, a gente sabe muito rápido, o que funcionou ou não. Pro nosso local é assim que funciona. Pro nacional a gente trabalha com o sistema de oferta, então, a gente oferece um conteúdo para a home do ge.globo, ou para ge.globonacional, aí eles escolhem, então depende da nossa habilidade de construir matérias com relevância nacional, aí pode ser um personagem aqui que o Vasco tem interesse em contratar, isso vai para a página nacional, aí a gente tem muito acesso da torcida do Vasco, e é automático, entrou no nacional a audiência cresce demais, cresce muito. Nossa missão é vender, convencer eles a utilizar. Por exemplo, nesse momento aqui, a gente tá com uma sobre o Vitinho que é um atacante do Atlético sendo sondado pelo Fluminense e Vasco. Isso é receita básica que vai para home deles lá. Porque tem Fluminense e Vasco, tem jogador importante. Isso a gente emplaca e aumenta nossa audiência. E aumentou. Entrou 14h44 e a audiência aumentou instantaneamente. Já tem 3 pessoas lendo essa notícia nesse momento. 3 mil pessoas só nessa notícia. 12:52

g) Os usuários podem fazer comentários abertos nas matérias? (interatividade)

Comenta e entra, mas tem as diretrizes de ofensas, um filtro orgânico. É o algoritmo quem faz, pelo que a pessoa escreve. 13:17

h) Há uma preocupação com a busca da informação dentro da matéria? (memória) 14:17

Ela consegue fazer a busca, mas a busca que mais funciona normalmente é a busca segmentada. Porque a pessoa vai lá entra no portal, nacional ou nossa, ela vai lá em times, lá abra uma aba com todos os escudos dos times. Ela vai lá no Vasco e pá. Ou ela abre o ge.globo, geralmente nesse nicho, as pessoas que gostam, elas acessam, ali ela tem também a distribuição da home também. Agora tem mais

assunto de copa, naturalmente, mas você vai encontrar ali as questões com mais relevância. Por exemplo, esse que citei do Vitinho, que é uma notícia do blogue da Nadja, já tá lá na home do nacional também, o blogue da Nadja tá na home nacional, então um produto local tá no nacional por uma notícia que envolve Vasco e Fluminense, então é a habilidade da repórter para apurar, habilidade do time para vender, e aí a gente emplaca e sobe audiência, essa é a nossa receitinha.

i) Quem se preocupa com o designer do portal? (design homepage) 15:47

Toda edição, tudo é do repórter, hoje a gente não consegue nem trabalhar com uma pós edição, eles mesmos se organizam para escrever e o outro revisar, a gente não trabalha com editores, a gente não tem essa figura. Então, o repórter, ele escreve, seleciona fotos, seleciona o texto, precisamos de cada vez mais de profissionais completos, que façam a entrega do pacote completo, e façam ainda o processo de oferta, porque ele é muito dinâmico, ele não passa pela coordenação, por exemplo, ele é muito direto, o repórter tem um sistema que ele manda direto, já o .globo, tem uma série de processos, que ele funciona muito bem nisso. 16:44

j) Quem define quais multimídias serão utilizadas na matéria? (multimídia) 16:57

Tem um banco de site, e a gente trabalha também com as assessorias dos clubes que eles têm para utilizar.

k) Há alguma regra para que seja utilizado um, ou mais, link(s) de uma matéria já escrita e que seja relevante para o tema que está sendo escrito (hipertexto)? 17:32

Não, ele tem um formato que você tem ali título, subtítulo, chapéu ou não tem, se vai colocar na home, se tem a gravatinha. Tem como se fosse um manual de redação, o já.globo tem isso. Tem uma liberdade, mas dentro desse caderno que tem algumas determinações, mas é um padrão que se consegue trabalhar de várias formas.

l) A sugestão de inserção de algumas estruturas de navegação ou inserção de links para interação com o público da matéria, é definida por quem? Ou não existe? (interatividade) 18:18

Tem, consegue trabalhar com isso também. Por exemplo, num assunto da nossa home, nosso carrossel, você consegue colocar uma série de junto. Coloca esse do Vitinho e mais três assuntos relacionados ao mercado do Atlético. O repórter quem escolhe, ele tem que escolher e formatar isso. Ele tem essa ferramenta e pode escolher utilizar ou não.

7.1.3. Plataforma

a) Qual a importância do Google para o Jornalismo? 19:02

Primeiro tem uma questão do algoritmo, o google nos entrega muita informação do que a pessoa está procurando, isso nos ajuda a definir muito o conteúdo que vamos produzir. Por exemplo, as pessoas procuram muito, do nosso nicho, onde assistir tal partida. Então a gente coloca isso no nosso título, então, é uma informação que a pessoa procura, tá no nosso título e tá na nossa matéria, obviamente. Para nós

parece meio simples, mas hoje, com essa pluralidade dos direitos de transmissão, as pessoas procuram muito onde vai passar Atlético de Madrid da liga dos campeões. Então, o Google nos mostra muito isso, quais as últimas pesquisas na última meia hora. Ou o Galvão falou sobre o Bruno Guimarães ter começado no Atlético, isso vai na hora, pessoas pesquisam, a gente teve link antigo, que as pessoas procuraram, porque algo foi comentado no jogo do Brasil. Então, a gente trabalha assim com o Google.

b) É verificado o volume de pesquisa sobre a notícia que está sendo feita (volume de pesquisa da palavra-chave)? 20:20

Tem uma ferramenta específica do google, não é só pelo *Google Trends* e também com o analytics, que nos dá uma leitura pelo que tem acontecido naquele momento, que tipo de público tem consumido, qual a palavra tem buscada naquele momento, as cinco últimas palavras buscadas na última meia hora. Eu tenho um acesso mais amplo, mas o repórter também tem acesso a principalmente ao desempenho em tempo real. Para já ter a medida do que está rendendo ou não.

c) É utilizado algum CMS (Content Management System)? (plataforma de automatização de conteúdo) 22:04

Aqui é o que eles chamam de backstage, mais ou menos o formato do wordpress, mas é algo específico feito para a globo. Ele é automatizado igual.

d) Há interação entre o site da sua empresa e o público interagente, disponível em todas as páginas, como um botão de aplicativo de mensagens? (interface) 22:47

Não tem, o que a gente tem feito mais, e é um produto que a globo tem apostado muito é um formato de livespré e pós jogo, e isso rende interatividade pelo twitter, que é uma rede social que no futebol funciona muito, ela ainda entra como interatividade, mas ela ainda pode funcionar melhor.

e) Há formulário de cadastro para o público receber mais notícias parecidas? (interface) 23:57

Não tem newsletter, a globo não trabalha com esse formato de assinatura.

7.1.4. SEO / TagTitle

a) Você acha importante as regras SEO? 24:43

É difícil, por qual linha você pergunta? Eu acho que a gente não pode virar receita, a gente precisa utilizar como ferramenta, importante e significativa, mas se a gente fica obcecado por ela, a gente entra na armadilha da caça clique, de você gerar conteúdo sem profundidade, sem que respeite a tua identidade, teu modelo de negócio, a gente trabalha muito aqui no GRPCOM que isso proteja a percepção de imagem que pessoa terá do nosso produto, então, eu vejo como ferramenta mas sem ser refém.

b) Quais das regras são mais importantes na escrita das matérias do portal? 26:26

A gente cuida muito da questão do título, às vezes mudamos uma palavra que nem seja a mais bonita ali, mas ela funciona melhor, tem um rendimento melhor, sem ser refém, sem ferir, mas normalmente é isso o que a gente faz.

c) Há um especialista em SEO no portal? 27:06

Temos o Rodrigo Saviani, mas é um repórter que por desenvolvimento próprio, que tem um olhar mais amplo, e ajuda os outros até nesse aspecto.

d) Há alguma orientação para os jornalistas escreverem as matérias para o portal? 27:45

Só se os repórteres buscarem por ele.

e) A parte de códigos do portal, onde são lidas as tagtitles, metadescriptions, tem alguma interferência do jornalista? 28:18

É automático.

f) Você percebe alguma semelhança entre algumas regras do Manual de Redação e as regras SEO? 28:39

Eu diria que é uma versão atualizada que é o que temos hoje na sociedade, é uma revolução tecnológica, e hoje eu encaro sim como uma espécie de manual, mas dentro das ferramentas que os jornalistas trabalham, não só aquele manual de redação antigo, que ele não funcionaria hoje porque ele seria muito menos funcional perto das necessidades que a gente, então, ele seria uma versão atualizada mesmo.

7.2 VISÃO GERAL DE ANÁLISE DO ESTADO DA ARTE

Título	Relevância pelo título	Categoria
A Hybrid Model Ranking Search Result for ResearchPaperSearchingon Social Bookmarking	N	Social Bookmarking
TravelRecommendationUsingGeo-taggedPhotos in Social Media for Tourist	N	Mídia Social
Ananalysisofusers' behaviourpatterns in theorganisationofinformation: A case studyofCiteULike	N	Geografia
LocationExtractionfrom Social Media: Geoparsing, LocationDisambiguation, andGeotagging	N	Mídia Social
Exploitingthewisdomof social connections to make personalizedrecommendationsonscholarlyarticles	N	Social Bookmarking
DuplicateDetection ProgrammingQuestionAnsweringCommunities in	N	Mídia Social
Witnessing in the New MemoryEcology:	N	Mídia Social

Memory Construction of the Syrian Conflict on YouTube		
The Messages Presented in Electronic Cigarette-Related Social Media Promotions and Discussion: Scoping Review	N	Mídia Social
Visibilidad web en buscadores factores algorítmicos de SEO onpage (FAOP) como técnica e práctica periódica	N	Otimização de Resultados
A context-aware personalized travel recommendation system based on geotagged social media data mining	N	Mídia Social
Quantifying the digital traces of Hurricane Sandy on Flickr	N	Mídia Social
Identification of tweets that mention books	N	Mídia Social
Does tagging improve the navigation of online recorded lectures by students?	N	Social Bookmarking
Emotion classification using a CNN LSTM based model for smooth emotional synchronization of the humanoid robot REN XIN	N	Saúde
Exploring the Distribution Patterns of Flickr Photos	N	Mídia Social
Assistive tagging: A survey of multimedia tagging with human computer exploration	N	Social Bookmarking
Multi-modal Similarity Retrieval with Distributed Key-value Store	N	Imagem
Tags in the Catalogue: Insights From a Usability Study of Library Thing for Libraries	N	Social Bookmarking
Using social media to quantify spatial and temporal dynamics of nature-based recreational activities	N	Mídia Social
Does the tag "islamic" help in customer satisfaction in dual banking sector?	N	Bancário
Video Captioning with Multi-Faceted Attention	N	Mídia Social
Analyzing transfer learning impact in biomedical cross-lingual named entity recognition and normalization	N	Saúde
Nursing staff participation in end-of-life nutrition and hydration decision-making in a nursing home: A qualitative study	N	Saúde
Improving geographic information retrieval in spatial data infrastructures	N	Geografia
How Deep Features Have Improved Event Recognition in Multimedia: A Survey	N	Multimídia
A Critical Perspective on Mental Health News in Six European Countries: How Are "Mental Health/Illness" and "Mental Health Literacy" Rhetorically Constructed?	N	Saúde
Vision-based real estate price estimation	N	Imobiliário
Measuring Destination Image through Travel Reviews in Search Engines	N	Pesquisa
Measuring Gastronomic Image Online	N	Turismo
The impact of semantic annotation techniques on content-based video lecture recommendation	N	Educação

Quantitative Risk Assessment of Norovirus Transmission in Food Establishments: Evaluating the Impact of Intervention Strategies and Food Employee Behavior on the Risk Associated with Norovirus in Food	N	Saúde
Customized Root-Analogue Implants: A Review on Outcomes from Clinical Trials and Case Reports	N	Saúde
Development of image collection representations for intelligent distributed systems	N	Social Bookmarking
Effects of Curcumin on Glycemic Control and Lipid Profile in Polycystic Ovary Syndrome: Systematic Review with Meta-Analysis and Trial Sequential Analysis	N	Saúde
The Pre-Implantation Embryo Induces Uterine Inflammatory Reaction in Mice	N	Saúde
How Do I Refactor This? An Empirical Study on Refactoring Trends and Topics in Stack Overflow	N	Refatoração
A Smartphone-Based Crowd-Sourced Database for Environmental Noise Assessment	N	Ruídos do Ambiente
An incremental probabilistic model for temporal theme analysis of landmarks	N	Mídia Social
Ensemble of Deep Models for Event Recognition	N	Computação
Prioritization of Susceptibility Genes for Ectopic Pregnancy by Gene Network Analysis	N	Saúde
A heuristic approach on metadata recommendation for search engine optimization		Otimização de Resultados
Neural ParsCit: a deep learning-based reference string parser	N	Saúde
An Embedded-Based Weighted Feature Selection Algorithm for Classifying Web Document		Otimização de Resultados
Attenuation of the increase in intraocular pressure with dexmedetomidine: Systematic review with meta-analysis and trial sequential analysis	N	Saúde
Write the story you want to read: world-queering through slash fiction creation	N	LGBTQI+
A review on human action analysis in videos for retrieval applications	N	Vídeo
Systematic Analysis of the Molecular Mechanism Underlying Decidualization Using a Text Mining Approach	N	Saúde
On the Prediction of Flickr Image Popularity by Analyzing Heterogeneous Social Sensory Data	N	Mídia Social
Extracting insights from big social data for smarter tourism destination management	N	Turismo
Tiered Governance and Demonetization: The Shifting Terms of Labor and Compensation in the Platform Economy	N	Economia
Group cognitive-behavioural therapy for hoarding disorder: Systematic review and meta-analysis	N	Saúde

Whatevidenceexists for the impact of climate change on the physiology and behaviour of important aquaculture marine crustacean species in Asia? A systematic map protocol	N	Saúde
Assessing similarity matching for possible integration of feature classifications of geospatial data from official and informal sources	N	Geografia
Mapping the field: a bibliometric analysis of land use and carbon emissions (LUCE) research from 1987 to 2018	N	Geografia
Building socially-enabled event-enriched maps	N	Geografia
Quantifying Tourist Behavior Patterns by Travel Motifs and Geo-Tagged Photos from Flickr	N	Mídia Social
Is there anything outside the tags?	N	Social Bookmarking
Digital chemsex publics: Algorithmic and user configurations of fear and desire on Pornhub	N	Social Bookmarking
Using Social Media to Discover Public Values, Interests, and Perceptions about Cattle Grazing on Park Lands	N	Mídia Social
Hierarchical Contaminated Web Page Classification Based on Meta Tag Denoising Disposal	N	Otimização de Resultados
Information search by applying VDL-based icon tags: an experimental study	N	Design
Exploring Land Use and Land Cover of Geotagged Social-Sensing Images Using Naive Bayes Classifier	N	Geografia
Social Event Classification via Boosted Multimodal Supervised Latent Dirichlet Allocation	N	Mídia Social
From the Stage to the Ring: The Early Years of Japanese Womens Professional Wrestling, 1948-1956	N	Esporte
Whatevidenceexists on how changes in marine ecosystem structure and functioning affect ecosystem services delivery? A systematic map protocol	N	Geografia
Transfer Learning of a Deep Learning Model for Exploring Tourist-Urban Image Using Geotagged Photos	N	Geografia
Dual Structure Constrained Multimodal Feature Coding for Social Event Detection from Flickr Data	N	Mídia Social
Wearable Wireless Biosensor Technology for Monitoring Cattle: A Review	N	Geografia
Development of Knowledge Base Using Human Experience Semantic Network for Instructive Texts	N	Organizacional
Outcomes of wildlife translocations in protected areas: what is the type and extent of existing evidence? A systematic map protocol	N	Geografia

7.3 VISÃO GERAL DE ANÁLISE DO ESTUDO DE CASO

Da ta	Título H1	Slug	TagTitle	TagDescription	Subtítulo	TÍ TU LO =	SUB TÍ TU LO = TAG	TÍ TU LO = TAG TITL	S L U G
-------	-----------	------	----------	----------------	-----------	------------	--------------------	---------------------	---------

						SL UG	DES C	E	= TAG TIT LE
02/ 11	Já campeão, Palmeiras recebe o Fortaleza em clima de festa; veja live	https://ge.globo.com/futebol/video/aos-30-min-Do-2o-tempo-gol-de-cabeca-de-ale-do-america-mg-contra-o-internacional-11090903.ghtml	</title><meta name="title" content="Aos 30 min do 2º tempo - gol de cabeça de Alê do América-MG contra o Internacional">	<meta name="description" content="Aos 30 min do 2º tempo - gol de cabeça de Alê do América-MG contra o Internacional">	Veja o gol do América-MG que garantiu o título do Palmeiras				X
	Palmeiras fatura cerca de R\$ 45 milhões com premiação por título e terá extra de patrocinadora	https://ge.globo.com/futebol/times/palmeiras/noticia/2022/11/02/palmeiras-fatura-cerca-de-r-45-milhoes-com-premiacao-por-titulo-e-tera-extra-de-patrocinadora.ghtml	</title><meta name="title" content="Palmeiras fatura cerca de R\$ 45 milhões com premiação por título e terá extra de patrocinadora">	<meta name="description" content="Prêmio pago pela CBF ao campeão é o maior da história da competição nacional">	Prêmio pago pela CBF ao campeão é o maior da história da competição nacional	X	X	X	X
	Abel iguala Felipão	https://ge.globo.com/futebol/times/palmeiras/noticia/2022/11/02/abel-igual-a-felipao-na-lista-de-tecnicos-com-mais-titulos-no-palmeiras.ghtml	</title><meta name="title" content="Abel iguala Felipão na lista de	<meta name="description" content="Português chega a	Português chega a seis taças		X		X

	na lista de técnicos com mais títulos no clube		técnicos com mais títulos no Palmeiras">	seis taças com a conquista do Brasileirão de 2022; veja os números">	com a conquista do Brasileirão de 2022; veja os números				
VÍDEOS	Após boa troca de passes, Churín empatou para o Dragão	https://ge.globo.com/futebol/video/aos-28-min-do-1o-tempo-gol-de-dentro-da-area-de-churin-do-atletico-go-contr-o-santos-11091154.ghtml	</title><meta name="title" content="Aos 28 min do 1º tempo - gol de dentro da área de Churín do Atlético-GO contra o Santos">	<meta name="description" content="Aos 28 min do 1º tempo - gol de dentro da área de Churín do Atlético-GO contra o Santos">	Aos 28 min do 1º tempo - gol de dentro da área de Churín do Atlético-GO contra o Santos		X		X
	Com revisão do VAR, Marcos Leonardo abre o placar para o Santos	https://ge.globo.com/futebol/video/aos-19-min-do-1o-tempo-revisao-do-var-de-marcos-leonardo-do-santos-contr-o-atletico-go-11091125.ghtml	</title><meta name="title" content="Aos 19 min do 1º tempo - revisão do VAR de Marcos Leonardo do Santos contra o Atlético-GO">	<meta name="description" content="Aos 19 min do 1º tempo - revisão do VAR de Marcos Leonardo do Santos contra o Atlético-GO">	Aos 19 min do 1º tempo - revisão do VAR de Marcos Leonardo do Santos contra o Atlético-GO		X		X

	Mano vê gosto amargo em derrota do Inter, mas promete luta pelo vice do Brasileiro	https://ge.globo.com/rs/futebol/times/internacional/noticia/2022/11/02/mano-ve-gosto-amargo-em-derrota-do-inter-mas-promete-luta-pelo-vice-do-brasileirao.ghtml	</title><meta name="title" content="Mano vê gosto amargo em derrota do Inter, mas promete luta pelo vice do Brasileiro">	<meta name="description" content="Treinador colorado valoriza campanha no retorno para confiar que grupo siga forte nas três rodadas restantes">	Treinador colorado valoriza campanha no retorno para confiar que grupo siga forte nas três rodadas restantes	X	X	X	X
	Derrota safada dessa do Inter nacional", analisa Luka	https://ge.globo.com/futebol/voz-da-torcida/video/derrota-safada-essa-do-internacional-analisa-luka-a-voz-da-torcida-11091131.ghtml	</title><meta name="title" content="" Derrota safada essa do Internacional"; analisa Luka A Voz da Torcida">	<meta name="description" content="" Derrota safada essa do Internacional"; analisa Luka A Voz da Torcida">	"Derrota safada essa do Internacional", analisa Luka A Voz da Torcida		aspas ão reconhecidas na tag	aspas ão reconhecidas na tag	X
	Mancini elogia atuação do América-MG; vitória dá fôlego	https://ge.globo.com/futebol/times/america-mg/noticia/2022/11/02/mancini-elogia-atuacao-do-america-mg-e-vitoria-da-folego-na-busca-pela-libertadores.ghtml	</title><meta name="title" content="Mancini elogia atuação do América-MG; vitória dá fôlego na busca pela Libertadores">	<meta name="description" content="Clube derrotou o vice-líder Internacional em casa e quebrou sequência de três partidas sem venceu em casa">	Clube derrotou o vice-líder Internacional em casa e quebrou sequência	X	X	X	X

	o na busc a pela Liber tador es				ncia de três partid as sem vence u em casa				
	PSG vence a Juve e e ence rra fase de grup o invict o	https://ge.globo.com/futebol/futebol-internacional/liga-dos-campeoes/jogo/02-11-2022/juventusita-parissaintgermain.ghtml	</title><meta name="title" content="Jogo Juventus 1 x 2 Paris Saint- Germain melhores momentos - Liga dos Campeões, Rodada 6 - tempo real ge">	<meta name="descripti on" content="Saiba tudo sobre o jogo Juventus 1 x 2 Paris Saint- Germain (Liga dos Campeões, Rodada 6): placar, escalação, fotos, gols e lances da partida.">	"Saiba tudo sobre o jogo Juven tus 1 x 2 Paris Saint- Germ ain (Liga dos Camp eões, Roda da 6): placar , escala ção, fotos, gols e lances da partid a."		X		
04/ 11	Rum o ao R\$ 1 bi de receit a, Flam engo arrec adou só com bilhet eria e	https://ge.globo.com/futebol/times/flamengo/noticia/2022/11/04/rumo-ao-r-1-bi-de-receita-flamengo-arrecadou-so-com-bilheteria-e-socio-torcedor-r-1439-mi.ghtml	</title><meta name="title" content="Rumo ao R\$ 1 bi de receita, Flamengo arrecadou só com bilheteria e sócios R\$ 143,9 mi">	<meta name="descripti on" content="Impuls ionado pelos bons resultados em campo, clubes alcançará sua projeção de receita bruta em 2022. Alto investimento no elenco fez a "conta a pagar"	Impul siona do pelos bons result ados em camp o, clubes alcanç ará sua projeç	X	X	X	X

sócios R\$ 143,9 mi			aumentar em R\$ 74 mi no último trimestre	ção de receita bruta em 2022. Alto investimento no elenco fez a "conta a pagar"; aumentar em R\$ 74 mi no último trimestre				
Corinthians tenta trocá-lo e no ataque, Santos se defende.	https://ge.globo.com/sp/futebol/jogo/04-11-2022/corinthians-santos.ghtml	</title><meta name="title" content="Confirma o jogo Corinthians 0 x 2 Santos - Campeonato Paulista Sub-20, Final - tempo real ge">	<meta name="description" content="Acompanhe o jogo Corinthians 0 x 2 Santos (Campeonato Paulista Sub-20, Final): escalação, fotos, vídeos, placar e lances da partida em tempo real.	Peixe abre placar com gol (quase) contra , amplia com Ivonei , de pênalti, e agora pode perder por até um gol de diferença na Vila				

	<p>Às vésperas da convocação Tite revela insônia: "Acorde as 4h"</p>	<p>https://g1.globo.com/futebol/selecao-brasileira/noticia/2022/11/04/a-dois-dias-da-convocacao-tite-revela-insonia-e-importancia-da-filha-na-carreira.ghtml</p>	<p></title><meta name="title" content="Às vésperas da convocação, Tite revela insônia e importância da filha na carreira"></p>	<p><meta name="description" content="Com bom humor, treinador da Seleção revela ao Esporte Espetacular que a esposa também tem sofrido antes da divulgação da lista para Copa, que acontece na próxima segunda-feira"></p>	<p>Com bom humor, treinador da Seleção revela ao Esporte Espetacular que a esposa também tem sofrido antes da divulgação da lista para Copa, que acontece na próxima segunda-feira</p>		X		
	<p>Rebaixado, Náutico se despede da Série B contra a Ponte Preta</p>	<p>https://g1.globo.com/pe/futebol/brasileirao-serie-b/jogo/04-11-2022/nautico-ponte-preta.ghtml</p>	<p></title><meta name="title" content="Assista ao jogo Náutico 0 x 0 Ponte Preta ao vivo - Campeonato Brasileiro Série B, Rodada 38 - tempo real g1"></p>	<p><meta name="description" content="Acompanhe o jogo Náutico 0 x 0 Ponte Preta (Campeonato Brasileiro Série B, Rodada 38). Veja no g1 os melhores momentos e tudo que você</p>	<p>Em partida fracamente, a Macaca derrotou o Timbuçá</p>				

				precisa saber em tempo real.	um gol de pênalti marcado por Léo Naldino no segundo tempo				
Superliga feminina de vôlei: Praia Clube x Sesc Flamengo	https://ge.globo.com/volei/ao-vivo/superliga-feminina-de-volei-praia-clube-x-sesc-flamengo.ghtml		</title><meta name="title" content="Superliga feminina de vôlei: Praia Clube 10 x 09 Sesc-Flamengo">	<meta name="description" content="Siga o clássico nesta sexta-feira; sportv2 transmite"	Siga o clássico nesta sexta-feira; sportv2 transmite	X	X		
Messi, Neymar e Mbappé: saiba os preços do trio no Cartola Express da Copa do Mundo	https://ge.globo.com/cartola-express/noticia/2022/11/04/messi-neymar-e-mbappe-saiba-os-precos-do-trio-no-cartola-express-da-copa-do-mundo.ghtml		</title><meta name="title" content="Messi, Neymar e Mbappé: saiba os preços do trio no Cartola Express da Copa do Mundo">	<meta name="description" content="Compãheiros do PSG vão custar C\$ 25, C\$ 24 e C\$ 23, respectivamente; você terá C\$ 160 para escalar os seus times no fantasy"	Compãheiros do PSG vão custar C\$ 25, C\$ 24 e C\$ 23, respectivamente; você terá C\$ 160 para escalar os seus times no fantasy	X	X	X	X

	Flamengo vence fora de casa, elimina o Ceará e se garante na final	https://ge.globo.com/pi/futebol/jogo/04-11-2022/ceara-flamengo.ghtml	</title><meta name="title" content="Jogo Ceará 0 x 2 Flamengo melhores momentos - Copa do Brasil Sub-20, Semifinal - tempo real ge">	<meta name="description" content="Saiba tudo sobre o jogo Ceará 0 x 2 Flamengo (Copa do Brasil Sub-20, Semifinal). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida."	Saiba tudo sobre o jogo Ceará 0 x 2 Flamengo (Copa do Brasil Sub-20, Semifinal). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida		X		
	Jogadora sub-17 da Ferroviária morre após sentir mal súbito durante treino	https://ge.globo.com/sp/ribeirao-preto-e-regiao/futebol/times/ferroviaria/noticia/2022/11/04/atleta-do-sub-17-da-ferroviaria-morre-apos-sentir-mal-subito-durante-treino.ghtml	</title><meta name="title" content="Atleta do sub-17 da Ferroviária morre após sentir mal súbito durante treino">	<meta name="description" content="Esther Martins Santos estava desde agosto na equipe e, de acordo com a Ferrinha, passou por todos os exames médicos sem apresentar anomalias">	Esther Martins Santos estava desde agosto na equipe e, de acordo com a Ferrinha, passou por todos		X		X

					os exames médicos sem apresentar anomalias				
	Copa do Mundo 2022 : Messi tem "chuteira de ouro" vazada; fotos	https://ge.globo.com/futebol/copa-do-mundo/noticia/2022/11/04/copa-do-mundo-2022-messi-tem-chuteira-de-ouro-vazada-fotos.ghtml	</title><meta name="title" content="Copa do Mundo 2022: Messi "chuteira de ouro" vazada; fotos">	<meta name="description" content="Site especializado vaza imagens da chuteira que o craque da Argentina irá utilizar na Copa no Catar"	Site especializado vaza imagens da chuteira que o craque da Argentina irá utilizar na Copa no Catar	X	X	X	X
	Rebeca alcança patamar de gênio no Mundial: "Vejo o quanto evolui"	https://ge.globo.com/ginastica/artistica/noticia/2022/11/04/rebeca-alcanca-patamar-de-genio-no-mundial-vejo-o-quanto-evolui.ghtml	</title><meta name="title" content="Rebeca alcança patamar de gênio no Mundial: "Vejo o quanto evolui":>	<meta name="description" content="Campêã olímpica impressiona por maturidade e profissionalismo para lidar com posto de maior estrela em Liverpool e confirmar título de número 1 do mundo">	Campêã olímpica impressiona por maturidade e profissionalismo para lidar com posto de maior estrela em	X	X	X	X

					Liverpool e confirmar título de número 1 do mundo				
Pedra no sapato: relembre 10 vezes que o Ituano derrubou gigantes do futebol brasileiro	https://ge.globo.com/sp/tem-esporte/futebol/times/ituano/noticia/2022/11/04/pedra-no-sapato-relembre-10-vezes-que-o-ituano-derrubou-gigantes-do-futebol-brasileiro.ghtml		<title><meta name="title" content="Pedra no sapato: relembre 10 vezes que o Ituano derrubou gigantes do futebol brasileiro">	<meta name="description" content="Galo de Itu enfrenta o Vasco na "final do acesso";, neste domingo, buscando surpreender um grande novamente"	Galo de Itu enfrenta o Vasco na "final do acesso", neste domingo, buscando surpreender um grande novamente	X	X	X	X
Daniel Alves revela cuidados para estar na Copa e elogia Neymar: "Gênio"	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/04/daniel-alves-revela-cuidados-para-estar-na-copa-e-elogia-neymar-genio.ghtml		</title><meta name="title" content="Daniel Alves revela cuidados para estar na Copa e elogia Neymar: "Gênio"">	<meta name="description" content="Sem jogos pelo Pumas-MEX, lateral treina no Barcelona B para estar em forma e ser lembrado por Tite. Jogador afirma que é quase impossível parar o camisa 10 da Seleção">	"Sem jogos pelo Pumas-MEX, lateral treina no Barcelona B para estar em forma e ser lembrado por	X	X	X	X

					Tite. Jogador afirma que é quase impossível parar o camisa 10 da Seleção				
	STJD marca para quarta o julgamento do caso entre Sport e Vasco	https://ge.globo.com/futebol/times/vasco/noticia/2022/11/04/stjd-marca-para-dia-9-11-julgamento-do-caso-entre-sport-x-vasco.ghtml	</title><meta name="title" content="STJD marca para próxima quarta julgamento do caso entre Sport x Vasco">	<meta name="description" content="Previsto para última quinta-feira, julgamento acontecerá no dia 9 de novembro e pode ser decisivo na luta pelo acesso à Série A">	Previsto para última quinta-feira, julgamento acontecerá no dia 9 de novembro e pode ser decisivo na luta pelo acesso à Série A		X		
6/11	Ja campeão, Palmeiras visita Cuiabá, que luta	https://ge.globo.com/mt/futebol/brasileirao-serie-a/jogo/06-11-2022/cuiaba-palmeiras.ghtml	</title><meta name="title" content="Assista ao jogo Cuiabá 1 x 0 Palmeiras ao vivo grátis - Campeonato Brasileiro, Rodada 36 - tempo real	<meta name="description" content="Acompanhe o jogo Cuiabá 1 x 0 Palmeiras (Campeonato Brasileiro, Rodada 36). Veja no ge os	Acompanhe o jogo Cuiabá 1 x 0 Palmeiras (Campeonato		X		

<p>contra degolada</p>		<p>ge"></p>	<p>melhores momentos e tudo que você precisa saber em tempo real."></p>	<p>Brasil eiro, Rodada 36). Veja no geos melhores momentos e tudo que você precisa saber em tempo real.</p>				
<p>De pênalti, Nenê abre o placar para o Vasco em Itu</p>	<p>https://ge.globo.com/futebol/video/aos-5-min-do-1o-tempo-gol-de-penalti-de-nene-do-vasco-contr-o-ituano-11101640.ghtml</p>	<p></title><meta name="title" content="Aos 5 min do 1º tempo - gol de pênalti de Nenê do Vasco contra o Ituano"></p>	<p><meta name="description" content="Aos 5 min do 1º tempo - gol de pênalti de Nenê do Vasco contra o Ituano"</p>	<p>Aos 5 min do 1º tempo - gol de pênalti de Nenê do Vasco contra o Ituano</p>				<p>X</p>
<p>De olho em milagre para subir, Sport encara Vila Nova</p>	<p>https://ge.globo.com/go/futebol/brasileirao-serie-b/jogo/06-11-2022/vila-nova-sport.ghtml</p>	<p></title><meta name="title" content="Jogo Vila Nova 0 x 0 Sport melhores momentos - Campeonato Brasileiro Série B, Rodada 38 - tempo real ge"></p>	<p><meta name="description" content="Saiba tudo sobre o jogo Vila Nova 0 x 0 Sport (Campeonato Brasileiro Série B, Rodada 38). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida."</p>	<p>Saiba tudo sobre o jogo Vila Nova 0 x 0 Sport (Campeonato Brasileiro Série B, Rodada</p>		<p>X</p>		

					da 38). Veja no ge os melho res mome ntos e o placar do jogo, gols e lances da partid a.				
	Torci das de Ituan o e Vasco brigam em frent e ao estád io antes de jogo decis ivo	https://ge.globo.com/sp/tem-esporte/futebol/brasileirao-serie-b/noticia/2022/11/06/video-torcidas-de-ituano-e-vasco-brigam-em-frente-ao-estadio-antes-de-jogo-decisivo-na-serie-b.ghtml	</title><meta name="title" content="VÍDE O: torcidas de Ituano e Vasco brigam em frente ao estádio antes de jogo decisivo na Série B">	<meta name="descripti on" content="Confu são nas imediações do Novelli Jr. é contida pela PM; equip es fazem jogo que vale acess o à elite nacio nal	Confu são nas imedi ações do Novelli Jr. é contida pela PM; equip es fazem jogo que vale acess o à elite nacio nal	X	X	X	X
	Philip pe Coutinho se machuca e está fora da Copa	https://ge.globo.com/futebol/brasil/selecao-brasil/noticia/2022/11/06/philippe-coutinho-se-machuca-e-deve-ficar-fora-da-copa-do-mundo.ghtml	</title><meta name="title" content="Philip pe Coutinho se machuca e está fora da Copa do Mundo">	<meta name="descripti on" content="Meia sofre lesão muscular em treino e não será convocado pela seleção brasileira"	Meia sofre lesão muscular em treino e não será convo cado	X	X	X	X

do Mundo				pela seleção brasileira				
Para não cair, Novorizontino visita já rebaixado Operário-PR	https://ge.globo.com/pr/futebol/brasileirao-serie-b/jogo/06-11-2022/operario-pr-novorizontino.ghtml	</title><meta name="title" content="Jogo Operário-PR 0 x 3 Novorizontino melhores momentos - Campeonato Brasileiro Série B, Rodada 38 - tempo real ge">	<meta name="description" content="Saiba tudo sobre o jogo Operário-PR 0 x 3 Novorizontino (Campeonato Brasileiro Série B, Rodada 38). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida.	Saiba tudo sobre o jogo Operário-PR 0 x 3 Novorizontino (Campeonato Brasileiro Série B, Rodada 38). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida.		X		
Raysa Leal é campeã da Liga Mundial	https://ge.globo.com/skate/noticia/2022/11/06/rayssa-leal-e-campea-da-liga-mundial-de-skate-street.ghtml	</title><meta name="title" content="Rayssa Leal é campeã da Liga Mundial de Skate Street">	<meta name="description" content="Brasileira venceu de virada na última manobra e corou uma temporada perfeita"	Brasileira venceu de virada na última manobra e corou	X	X	X	X

de Skate Street					u uma temporada perfeita				
Letícia Bufoni faz manobra inédita de skate durante salto de avião	https://ge.globo.com/skate/noticia/2022/11/06/leticia-bufoni-faz-manobra-inedita-de-skate-durante-salto-de-aviao.ghtml	</title><meta name="title" content="Letícia Bufoni faz manobra inédita de skate durante salto de avião">	<meta name="description" content="Brasileira é primeira mulher no mundo a andar numa pista a quase 9 mil pés de altura e acertar um feeble grind batizado de "skygrind"; skatista quer homologar feito no Guinness Book"	Brasileira é primeira mulher no mundo a andar numa pista a quase 9 mil pés de altura e acertar um feeble grind batizado de "skygrind"; skatista quer homologar feito no Guinness Book	X	X	X	X	
Neymar brilha com gol e assistência, e líder PSG vence	https://ge.globo.com/futebol/futebol-internacional/futebol-frances/jogo/06-11-2022/orient-parissaintgermain.ghtml	</title><meta name="title" content="Jogo Lorient 1 x 2 Paris Saint-Germain melhores momentos - Campeonato Francês, Rodada 14 -	<meta name="description" content="Saiba tudo sobre o jogo Lorient 1 x 2 Paris Saint-Germain (Campeonato Francês, Rodada 14).	Saiba tudo sobre o jogo Lorient 1 x 2 Paris Saint-Germain (Cam				X	

	e mais uma		tempo real ge">	Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida.	peona to Franc ês, Roda da 14). Veja no ge os melho res mome ntos e o placar do jogo, gols e lances da partid a.				
	Rebe ca é bron ze no solo na desp edida de Baile de Favela no Mun dial	https://ge.globo.com/ginastica-artistica/noticia/2022/11/06/rebeca-e-bronze-no-solo-na-despedida-de-baile-de-favela-no-mundial.ghtml	</title><meta name="title" content="Rebeca é bronze no solo na despedida de Baile de Favela no Mundial">	<meta name="description" content="Camp eã mundial do individual geral, Rebeca Andrade garante a medalha de bronze no solo em sua última apresentação com a música Baile de Favela; Arthur Nory também é terceiro na barra fixa"	Camp eã mundial do individual geral, Rebeca Andrade garante a medalha de bronze no solo em sua última apresentação com a música Baile de	X	X	X	X

					Favela; Arthur Nory também é terceiro na barra fixa				
	Sala marcada duas vezes, e Liverpool venceu o Tottenham em Londres	https://ge.globo.com/futebol/futebol-internacional/futebol-ingles/jogo/06-11-2022/Tottenham-Liverpool.ghtml	</title><meta name="title" content="Jogo Tottenham 1 x 2 Liverpool melhores momentos - Campeonato Inglês, Rodada 15 - tempo real ge">	<meta name="description" content="Saiba tudo sobre o jogo Tottenham 1 x 2 Liverpool (Campeonato Inglês, Rodada 15). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida.	Saiba tudo sobre o jogo Tottenham 1 x 2 Liverpool (Campeonato Inglês, Rodada 15). Veja no ge os melhores momentos e o placar do jogo, gols e lances da partida.		X		
7/1 x 1 BOT	Siga em tempo real - CAM	https://ge.globo.com/mg/futebol/brasileirao-serie-a/jogo/07-11-2022/atletico-mg-botafogo.ghtml	</title><meta name="title" content="Assista ao jogo Atlético-MG x Botafogo ao vivo - Campeonato	<meta name="description" content="Acompanhe o jogo Atlético-MG x Botafogo (Campeonato	Saiba tudo sobre o jogo Atlético-MG x Botafogo				

			Brasileiro, Rodada 36 - tempo real ge">	Brasileiro, Rodada 36). Veja no ge os melhores momentos e tudo que você precisa saber em tempo real."	go (Campeonato Brasileiro, Rodada 36). Veja no ge os melhores momentos e tudo que você precisa saber em tempo real.				
	Os caras do Brasil	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/convocados-da-selecao-para-a-copa-do-mundo-2022-veja-a-lista-de-tite.ghtml	</title><meta name="title" content="Convocados da Seleção para a Copa do Mundo 2022: veja a lista completa de Tite">	<meta name="description" content="Conheça os 26 jogadores que vão representar o Brasil no Mundial do Catar"	Conheça os 26 jogadores que vão representar o Brasil no Mundial do Catar		X	X	X
Firmino, Gabigol... veja nomes importantes que		https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/convocacao-da-selecao-veja-quem-ficou-fora-da-lista-final-de-tite.ghtml	</title><meta name="title" content="Convocação da Seleção: veja quem ficou fora da lista final de Tite">	<meta name="description" content="Técnico Tite anunciou os 26 nomes que vão formar o elenco da Seleção Brasileira no Catar e	Técnico Tite anunciou os 26 nomes que vão formar o		X		X

ficaram fora			algumas ausências chamaram atenção"	elenco da Seleção Brasileira no Catar e algumas ausências chamaram atenção				
Confirma a numeração que os jogadores da seleção usarão no Catar	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/veja-a-numeracao-dos-jogadores-da-selecao-brasileira-na-copa-do-mundo-de-2022.ghtml		</title><meta name="title" content="Veja a numeração dos jogadores da seleção brasileira na Copa do Mundo de 2022">	<meta name="description" content="Lista de convocados foi divulgada nesta segunda-feira; CBF diz que há diferenças"		X		X
Campeão mundial aos 17, Ronaldo diz que levaria Endrick para a Copa	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/ronaldo-fenomeno-diz-que-levaria-endrick-do-palmeiras-para-a-copa-do-mundo.ghtml		</title><meta name="title" content="Ronaldo Fenômeno diz que levaria Endrick, do Palmeiras, para a Copa do Mundo">	<meta name="description" content="Convocado em 1994 com apenas 17 anos, ex-atacante diz que sente falta deste tipo de convocação na seleção brasileira e cita Kaká como exemplo em 2002"		X		X

					de convo cação na seleçã o brasil eira e cita Kaká como exem plo em 2002				
	Real madr id cai para o Rayo Valle cano e sofre prim eira derro ta no Espa nhol	https://ge.globo.com/futebol/futebol-internacional/futebol-espanhol/jogo/07-11-2022/rayo-vallecano-real-madrid.ghtml	</title><meta name="title" content="Jogo Rayo Vallec ano 3 x 2 Real Madrid melhores momentos - Campeonato Espanhol, Rodada 13 - tempo real ge">	<meta name="descripti on" content="Saiba tudo sobre o jogo Rayo Vallec ano 3 x 2 Real Madrid (Campeonato Espanhol, Rodada 13): placar, escalação, fotos, gols e lances da partida."	Saiba tudo sobre o jogo Rayo Vallec ano 3 x 2 Real Madri d (Cam peona to Espan hol, Roda da 13): placar , escla ção, fotos, gols e lances da partid a.		X		
	Aliss on exalt a suc esso dos	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/alisson-exalta-sucesso-dos-goleiros-brasileiros-as-vesperas-da-copa.ghtml	</title><meta name="title" content="Alisso n exalta sucesso dos goleiros brasileiros às	<meta name="descripti on" content="Titular da seleção brasileira diz que se inspira	Titular da seleçã o brasil eira diz	X	X	X	X

goleiros brasileiros às vésperas da Copa		vésperas da Copa">	no irmão Muriel e em Taffarel, preparador de goleiros do Brasil e do Liverpool, e relembra história com o pai na Copa de 1998"	que se inspira no irmão Muriel e em Taffarel, preparador de goleiros do Brasil e do Liverpool, e relembra história com o pai na Copa de 1998				
Daniel Alves celebra convocação: "Dia de vitória"	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/daniel-alves-comemora-convocacao-para-a-copa-do-mundo-dia-de-vitoria.ghtml	</title><meta name="title" content="Daniel Alves comemora convocação para a Copa do Mundo: "Dia de vitória":>	<meta name="description" content="Em vídeo, lateral resalta superação depois ter ficado fora de 2018, e refuta críticas"	Em vídeo, lateral resalta superação depois ter ficado fora de 2018, e refuta críticas	X	X	X	X
Pedro comemora a convocação	https://ge.globo.com/futebol/times/flamengo/noticia/2022/11/07/pedro-do-flamengo-comemora-a-convocacao-para-copa-e-pede-a-namorada-em-casamento.ghtml	</title><metaname="title"content="Pedro, do Flamengo, comemora a convocação para Copa e	<meta name="description"content="Centroavante e Everton Ribeiro estão na lista do Brasil para o	Centroavante e Everton Ribeiro		X		X

ão e pede a namorada em casamento		pede a namorada em casamento">	Mundial do Catar"	estão na lista do Brasil para o Mundial do Catar				
Lista de Tite repercute na imprensa internacional: "Nov e ataca ntes" !	https://ge.globo.com/futebol/selecao-brasileira/noticia/2022/11/07/lista-de-tite-repercute-na-imprensa-internacional-nove-atacantes.ghtml	<title><meta name="title" content="Lista de Tite repercute na imprensa internacional: "Nove atacantes!"";>	<meta name="description" content="Jornal espanhol Marca destaca número de jogadores de ataque convocados. Ausência de Firmino também é manchete"	Jornal espanhol Marca destaca número de jogadores de ataque convocados. Ausência de Firmino também é manchete	X	X	X	X