



Construção de credibilidade dos dados: como dados auxiliares de SEO são validados

Building data credibility: How auxiliary SEO data is validated

Carlos Henrique Marcondes ^a 

Victor Martini Gabry ^{a, *} 

RESUMO: O presente trabalho buscou estudar de que forma instituições legitimadoras criam credibilidade para métricas em seus *campus* de atuação e influência. Para isso fizemos um recorte das formas como três companhias que negociam dados para o mercado de otimização mecanismos de busca – MOZ, SEMRush e Ahrefs – criavam discursos que validassem os indicadores que vendem. Tomamos como arcabouço teórico os *campus* de Bourdieu e as táticas e estratégias de Certeau, e como metodologia a análise do discurso de linha francesa de Algirdas Greimas. Observamos que as táticas usadas por instituições legitimadoras se valem da terceirização da responsabilidade, do ato de desmentir terceiros, da compartimentalização do dado e do solucionismo tecnológico como postulado por Morozov para ratificarem sua relevância. Também encontramos indícios de que a instalação de um novo “dado” por uma instituição legitimadora passa por um processo com certo grau de generalização de validação, em que o discurso leva em conta os atores envolvidos em seu uso e interpretação.

Palavras-chave: SEO. Otimização para Mecanismos de Busca. SEM. Marketing para Mecanismos de Busca. Dados.

ABSTRACT: This study aimed to explore how legitimizing institutions create credibility for metrics within their spheres of operation and influence. It specifically examined how three companies—MOZ, SEMRush, and Ahrefs—construct credible narratives to justify the indicators they sell, using a subset of search engine optimization metrics. The theoretical framework employed includes Bourdieu's field theory and Certeau's tactics and strategies, with the methodology being the French discourse analysis of Algirdas Greimas. We observed that legitimizing institutions rely on tactics such as outsourcing responsibility, denying the claims of others, compartmentalizing data, and embracing technological solutionism as postulated by Morozov to affirm their relevance. We also found evidence that the introduction of a new "data" by a legitimizing institution undergoes a process with a certain degree of generalized validation, where the discourse considers the actors involved in its use and interpretation.

Keywords: SEO. Search Engine Optimization. SEM. Search Engine Marketing. Data.

^a Programa de Pós-Graduação Ciência da Informação, Universidade Federal Fluminense, Niterói, RJ, Brasil.

* Correspondência para/Correspondence to: Victor Martini Gabry. E-mail: victorgabry@id.uff.br .

Recebido em/Received: 24/04/2024; Aprovado em/Approved: 10/10/2024.

Artigo publicado em acesso aberto sob licença [CC BY 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) 

INTRODUÇÃO

A corrida pela primeira posição nos mecanismos de busca revela muito mais do que uma simples disputa por atenção. Sendo o Google uma das principais formas de acesso à internet, atrás apenas da abertura de e-mails e do acesso às redes sociais de maneira geral (Dixon, 2023), seu impacto é sensível no processo de ganho informacional (Schultheiß, 2023).

O que determina os resultados dos mecanismos de busca pode ser interpretado mais como o produto de uma disputa entre táticas e estratégias (Certeau, 2014) do que como o resultado de um algoritmo neutro. Este artigo se insere como um esforço de problematizar os resultados trazidos pelos buscadores, orientados e direcionados por relações entre diferentes campos (Bourdieu, 1992), suas disputas e, o mais importante: as ferramentas de criação de credibilidade (Greimas, 2014).

A relevância da companhia Alphabet e da ferramenta Google, sobretudo no que tange a busca e recuperação de informação, é sentida em diversas áreas. Tanto antes quanto durante a pandemia do Coronavírus, por exemplo, o uso de ferramentas como o Google Trends (Nutti *et al.*, 2014) permitiu que se mapeasse as buscas realizadas por grandes volumes populacionais e sua relação com o começo de epidemias e crises sanitárias em diferentes partes do mundo. No âmbito político, a International Fact-Checking Network propôs ao consórcio Schema.org a criação de um código *schema* que permitisse às verificações de fatos uma posição de destaque nas páginas responsivas dos mecanismos de busca (SERP) (Gingras, 2016).

Todas essas aplicações – tanto científicas quanto mercadológicas – da otimização para mecanismos de busca (SEO, na sigla em inglês) por parte de desenvolvedores, pesquisadores e profissionais de mercado, demandam dados. Em nossa pesquisa pretendemos questionar quais dados são usados para essas atividades e como eles conseguem se alçar à uma posição de credibilidade entre os atores citados.

Em revisão sistemática da literatura (Gabry, Marcondes, 2024, no prelo) foram identificados dois consensos relevantes sobre o campo do SEO: primeiro, que há uma falta de embasamento teórico sobre os processos de otimização; segundo, que tanto a pesquisa científica quanto o mercado de trabalho se baseiam em dois perfis de dados para direcionar táticas visando indexação. Os primeiros são os dados oficiais do Google, através do Google Analytics, Google Search Console e ferramentas como o Google Trends e o Google AdWords. Os demais são fornecidos por ferramentas de terceiros, como os dados obtidos via *softwares* no modelo *martech*, a exemplo de SEMRush, Ahrefs, MOZ, AlsoAsked e Screaming Frog (Escandell-Poveda, Papí-Gálvez, Iglesias-García, 2023; Paul, Aithal, 2018; Sahin, Dirsehan, 2023).

O Google se coloca no fornecimento de dados sobre posicionamento e tráfego online em uma posição ambígua. Por um lado, há o interesse não dito da companhia de fornecer os melhores resultados, precisando, para isso, educar em parte os produtores de conteúdo para gerarem páginas “encontráveis” por seu mecanismo. Por outro, não é possível disponibilizar todos os dados que seriam relevantes para alguém que busque

um bom posicionamento, para não se incorrer no risco competitivo de deixar o código proprietário muito “óbvio” para a concorrência.

Concorrentes diretos de um domínio, posição absoluta que um website alcança em *queries* específicas, classificação das palavras-chave quanto à sua intenção de busca (transacional, informacional, navegacional (Broder, 2002)), não são dados fornecidos pelas ferramentas do Google, mas por companhias que concorrem indiretamente pela credibilidade quanto ao estado real das SERP. Aqui, agências e profissionais de SEO e SEM podem gastar de R\$ 600 até R\$ 15 mil mensalmente¹ para ter acesso a uma variedade de ferramentas em busca de vantagem competitiva.

Entretanto, entendendo que esses dados são proprietários do Google, há de se questionar como essas empresas têm acesso a essas informações a ponto de comercializá-las. Essa informação, valiosa tanto para a dinâmica de anúncios, principal origem do faturamento da Alphabet, quanto para os resultados orgânicos, essencial para as diferentes empresas que visam alto posicionamento (Kritzinger, 2017; Sen, 2005; Wang, Hao, 2011; Yeoung, Nisar, 2014) não estaria acessível a nenhuma ferramenta sem incorrer em quebra de informação sensível dos códigos do Google. Por isso se torna objetivo deste artigo entender como os vendedores de dados sobre o Google conseguem construir credibilidade para seus ativos.

De modo a descortinar a estratégia que é dada para atribuir credibilidade a esses dados, iremos nos valer da análise semiótica do discurso na abordagem do quadro greimasiano do conceito de veridicção (Greimas, 2014), expondo se e como as estratégias chamadas camuflagem subjetivante e objetivante são usadas. A análise semiótica do discurso se vale das construções de percursos gerativos de sentido entre os conceitos de ser, não-ser, parecer e não-parecer, no que se chama de quadro semiótico, para estabelecer o que a narrativa construída pelo enunciado atribui valor verdade (veridictório), de mentira, de segredo e de falsidade.

Como corpus serão analisadas as páginas explicando os conceitos de Authority Score (SEMRush)², Domain Authority (MOZ)³ e Domain Rating (Ahrefs)⁴, noções caras aos profissionais de SEO, uma vez que todos esses indicadores buscam replicar o Page Rank, métrica usada pelo algoritmo do Google para indicar relevância de uma página (Sergei, Brin, 1998). Esperamos, com isso, explorar detalhes sobre o exercício da prática profissional no campo do SEO, descortinar a economia envolvida em diferentes companhias, e como táticas e estratégias acontecem em torno da credibilidade de dados e sua comercialização.

¹ Dados de referência extraídos dos sites proprietários das referidas ferramentas, considerando-se os pacotes básicos aos mais caros de uma soma de diferentes ferramentas dessas.

² <https://pt.semrush.com/kb/747-authority-score-backlink-scores>

³ <https://moz.com/learn/seo/domain-authority>

⁴ <https://ahrefs.com/pt/seo/glossary/domain-rating>

DESENVOLVIMENTO

De modo a explicitar as formas como os dados dessas companhias são envernizados com matizes de credibilidade, o trabalho está dividido como segue: primeiro, descreveremos as estratégias greimasianas e o método da sua análise semiótica; depois, faremos a análise das páginas dedicadas a definir cada uma das métricas; por fim, discutiremos os achados à luz do arcabouço teórico proposto para o estudo do campo do SEO.

As estratégias de veridicção greimasianas

Na análise semiótica do discurso, há diferentes níveis de análise, cujo intercruzamento tornam os sentidos construídos pelo enunciado tangíveis para reflexão. Em particular a forma como um enunciado crível é investido desse sentido é objeto privilegiado de reflexão do fundador da linha da semiótica francesa, Greimas (2014). Greimas dedica o capítulo “O contrato de veridicção” a afirmar e comprovar que o conteúdo dito verdadeiro nada mais é do que um discurso adequado à sua janela de tempo e espaço, circunscrito em uma sociedade por uma forma. Ele identifica esta forma, nos discursos ocidentais contemporâneos, a partir de duas estruturas principais: a camuflagem subjetivante e a camuflagem objetivante.

Sobre a camuflagem subjetivante o autor a compara com a estrutura do segredo: para que algo seja considerado em um formato de “verdade”, deve tomar a forma de algo secreto, passado entre os iniciados (Greimas, 2014, p. 123).

O primeiro tipo de manipulação, da ordem do que se pode chamar camuflagem subjetivante, é bem ilustrado pelo discurso lacaniano que, como confessa o próprio autor, deve ser construído de maneira tal que, para ser aceito como ‘verdadeiro’, deve parecer ‘secreto’. O discurso que existe apenas para sugerir um plano anagógico a ser decifrado, avatar moderno do ‘discurso de parábolas’ de Jesus, é uma das formas da comunicação veridictória assumida. (GREIMAS, 2014, p. 123).

Essa mesma estrutura foi amplamente explorada por Rocha (2021) quando este traça todo o universo de referências internas do movimento bolsonarista. No livro fica claro como a narrativa em torno do “Orvil”, a ameaça secreta do golpe comunista e os inimigos secretos do regime são usados para criar uma narrativa veridictória para os adeptos visando cooptar sujeitos através do discurso. Essa é a forma da primeira camuflagem.

A segunda camuflagem é a objetivante, em uma estrutura que é aquela do fazer científico: onde o enunciador do discurso se omite em prol do enunciado, dando-lhe uma aparência de “ser natural”, orgânico, que se dá sozinho.

A essa comunicação hermético-hermenêutica se opõe o discurso científico – ou que se pretende como tal – este da ordem da camuflagem objetivante: para ser aceito como verdadeiro, ele procura parecer não o discurso de um sujeito,

mas o puro enunciado das relações necessárias entre as coisas, e para isso apaga, tanto quanto possível, todas as marcas da enunciação. Um enunciado como *a terra é redonda* certamente pressupõe, como sabemos, o *eu digo que... eu sei que... eu estou certo que... a terra é redonda*. Isso não impede que todo esse suporte enunciativo, que insere o enunciado no contexto de uma comunicação corriqueira, seja oculto para, a rigor, deixar transparecer apenas o impessoal “é verdade que...”, o qual se espera que o modalize, conferindo-lhe toda objetividade. (GREIMAS, 2014, p. 123).

Assim é o discurso científico, onde a primeira pessoa é abolida, e as relações entre os objetos se dão de maneira natural. É importante frisar que essa estrutura de credibilidade do discurso não é criticada nem isentada pelo autor, apenas explicitada, de modo a que narrativas que usem a mesma estrutura, sem o mesmo compromisso que a científica, não passem ilesas pelo engodo.

Ambas as camuflagens podem ser usadas nos discursos veridictórios para criar o sentido de credibilidade do discurso. Eles são mais vistos durante a fase de manipulação do texto enunciado, fase do nível narrativo de uma análise em que “um sujeito age sobre outro para levá-lo a querer e/ou dever fazer alguma coisa” (Fiorin, 2018, p. 29). Aqui as camuflagens tomam forma narrativa, na construção dos enunciados, sendo no nível discursivo que há uma manifestação dessa camuflagem palpável nas estratégias utilizadas pelo discurso.

Nosso objeto empírico – os dados oferecidos por companhias que chamamos de Instituições de Legitimação (IL) de acordo com a teoria – é sintomático do que Morozov (2018) chama de “solucionismo”, um movimento cultural iniciado no Vale do Silício, já difundido em políticas públicas, que assume que decisões mais informatizadas e baseadas em mais dados serão melhores. Essa preposição se confunde com a necessidade de obter dados, fazendo com que este esteja acima das soluções efetivamente ou dos impactos das soluções.

Em um objeto que oferece dados que não podem ser comprovadamente efetivos, não vão direto à fonte e apenas se propõem a ser os corretos, como convencer os leitores e clientes de sua efetividade? Essa pergunta, por uma lógica hipotético-dedutivo, nos ajuda a identificar como uma cultura de dados constrói a sua credibilidade.

Por isso, analisaremos o percurso gerativo de sentido, analisando como e de que forma as camuflagens surgem, de modo a evidenciar, de maneira mais abrangente, como se justifica a adesão a essas métricas.

Breve descrição do objeto

No modelo estipulado, as companhias analisadas são caracterizadas como instâncias legitimadoras (Bourdieu, 1992) de segundo grau, uma vez que não são formuladoras de estratégias (Certeau, 2014), mas sim de táticas. Usamos esse como um fator hierarquizante, apesar de Bourdieu, mesmo reconhecendo um grau de maior ou menor poder entre os sujeitos, não estruturou seu trabalho em termos de hierarquia. Elas são

assim classificadas pois seus dados exercem funções legitimadoras das táticas performadas por produtores culturais frente aos consumidores (seus clientes), mas não são de primeira importância: reservamos este lugar para os atores capazes de fazer estratégias.

Esses dados funcionam tentando replicar os métodos do Google através de um método dedutivo dos resultados das SERP, em um tipo de engenharia reversa. O objetivo das companhias objeto do estudo – SEMRush, Ahrefs, Moz – é vender para empresas terceiras informações privilegiadas sobre os seus rankings e os rankings de seus concorrentes. Para se diferenciarem umas das outras, suas métricas recebem nomes diferentes, são empregadas táticas de *branded content* e criam-se certificados para uso de suas ferramentas.

Sua estrutura de negócios se sustenta partindo do pressuposto de que os dados oferecidos são muito próximos dos resultados reais dos buscadores, sem, com isso, dizerem que “desvendaram o algoritmo” dos buscadores. Isso se dá por dois motivos: um legal, onde afirmar tal coisa poderia acarretar diferentes processos e investigações de práticas ilegais de competição; e um prático, onde ofertar o código certo e o dado definitivo implicaria que alguém teria o resultado de posicionar na frente dos demais indefinidamente, com um valor incalculável de mercado.

Tais empresas justificam seus indicadores com páginas dedicadas em seus websites a explicar como o dado é construído, o que pode ser feito com ele e como analisar as suas oscilações. Esse guia interpretativo é o que iremos analisar a seguir.

Figura 1. print do website Authority Score (SEMRush)



Authority Score (SEMRush)

O Authority Score é a métrica determinada pelo SEMRush em diferentes abas de seu software. A companhia assim a define:

Semrush Authority Score é nossa métrica composta usada para medir a qualidade e performance de SEO geral de domínios ou webpages. Se baseia em uma série de métricas que provam autoridade e confiabilidade. (Team, 2020, p. tradução do autor)

Mais adiante o texto indica que essa métrica pode ser encontrada em outras ferramentas do próprio software – essa definição, em particular, não usa a primeira pessoa que vemos em “nossa” na definição, começando a performance (Fiorin, 2018) de conjunção entre o sujeito “métrica” e o objeto “crível”. Esse percurso vai ser exercido, nesta métrica, através de jogos de camuflagem como esse: em que por vezes se usa a primeira pessoa, e por vezes a oculta.

Em um intertítulo “como o authority score é feito”, o artigo usa outra estratégia de camuflagem objetivante, quando afirma que “o Authority Score é o resultado de cálculos de IA e aprendizado de máquina para mensurar um domínio baseado em três aspectos”. Tal definição, que já se vale de dois outros sujeitos que ajudam na conjunção – a IA e o aprendizado de máquina – isenta a própria companhia de ter desenvolvido tanto a IA quanto o aprendizado de máquina (dois sujeitos que, à rigor metodológico, são a mesma coisa) que formulariam a métrica.

A composição de três facetas é uma tática discursiva de composição. Podemos traçar um paralelo entre essa composição do índice em facetas com a narrativa do monomito (Campbell, 2007), especificamente as provações.

O sujeito-dado precisa passar pela decomposição em três facetas para provar seu valor como um dado seguro e confiável. Além disso, cada uma dessas facetas é também sujeita à prova, subdivididas em seis aspectos, em uma narrativa que busca valorizar a complexidade, tratada como aspecto de uma credibilidade do indicador.

O artigo segue explicando como o Authority Score deve ser usado para comparações moderadas. Aqui o sujeito e sua conjunção com a credibilidade precisam se antecipar à objeção, esse percalço na jornada que ameaça a sanção positiva – o valor a ser validado ao final da leitura. Para isso a narrativa interpõe possíveis usos da métrica que podem ou não ser considerados como válidos, mostrando que o dado pode ser usado para comparações apenas entre competidores de mesmo porte.

Chega o momento, sob o intertítulo “Como usar o Authority Score para SEO” que a narrativa começa o momento de “retorno para casa” do herói transformado, mostrando como são as táticas possíveis com o uso dessa métrica. Aqui são novamente listados – como se a multiplicação de soluções endossasse pela composição – quatro aspectos onde a métrica faria a diferença: “análise da competição”, “avaliar propostas de link building”, “acompanhar o impacto de uma campanha de SEO” e “monitorar SEO negativo”.

É importante destacar que a forma como cada um desses percursos gerativos de sentido narrativo, dessas performances, tem um sujeito que entra em conjuntura com o dado para alcançar um objeto. Realizar a análise da competição, avaliar propostas, acompanhar campanhas, monitorar, todos esses atos visam uma outra narrativa, para

a qual sobe-se um nível de análise, para o nível fundamental, onde as categorias mais abstratas de “sucesso” e “lucro” só se tornam possíveis ao enunciador caso este aceite a credibilidade do dado em questão. Esse é um discurso que coaduna com aquele destacado por Morozov (2018) do solucionismo tecnológico, e que, como vemos, é o discurso que permeia todos os conteúdos de validação dos dados.

Dadas as performances possíveis que o sujeito munido dos dados pode fazer, o texto se encaminha para a última parte, onde ensina como fazer essa métrica crescer. Mas qual seria o interesse de dar um dado e ensiná-lo a fazê-lo crescer, se a lógica competitiva de um mecanismo de busca, onde apenas um “primeiro lugar” é possível resultaria em todos os clientes, menos um, da plataforma falhando em alcançar a sanção?

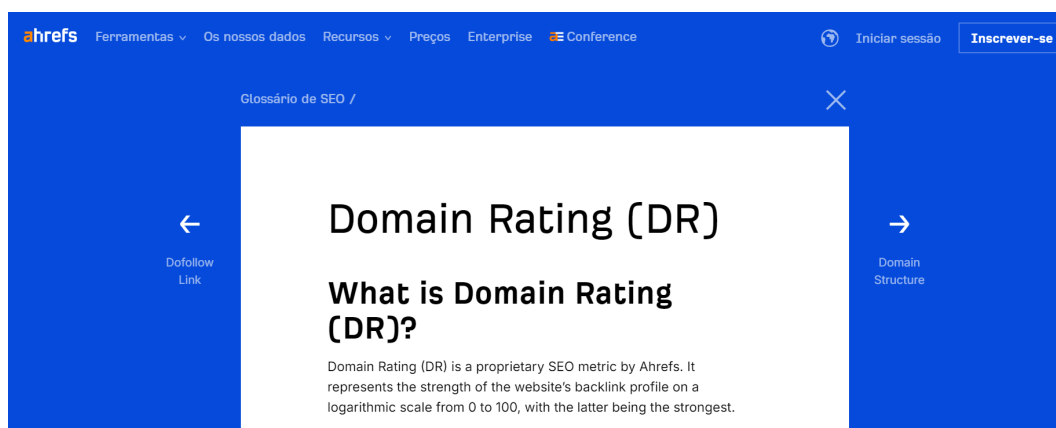
Esse ponto é chave para entender que o discurso construído não é voltado para o negócio que contrata o software ou a empresa que será cliente da agência que o contrata, mas que o dado serve para justificar os investimentos realizados em SEO pelo contratante. O texto, portanto, deixa como não-dito que a métrica é para profissionais de SEO orientarem seu trabalho e seu valor para outrem.

Neste ponto, novamente o argumento da compartimentação para criação de credibilidade aparece – dessa vez, porém, entre “como fazer o dado crescer”. Os três caminhos que o texto oferece para o leitor são: conseguir mais tráfego orgânico para o seu site; construir mais links de alto nível; diminuir o fator de toxicidade dos links. Há toda uma reflexão que podemos fazer entre as escolhas dos termos: “orgânico” e “tóxico” dialogam com outras isotopias para construir o universo discursivo do campo do SEO.

Enquanto “orgânico” faz referência ao tráfego adquirido por “vias naturais”, ou, definindo-o tecnicamente, pelas vias de mecanismos de busca em links não patrocinados, os links “tóxicos” são aqueles marcados pelos algoritmos como estigmatizados por se originarem de domínios com falhas de segurança modernas, como a não adoção de protocolos HTTPS, participantes de redes de backlink (quando diversos sites trocam links uns com os outros para construir uma “falsa autoridade” e então vender links para sites fora do grupo) ou com excesso de táticas de SEO Black Hat (as táticas que não concordam com o que os mecanismos de busca, em suas diretrizes oficiais, afirmam ser a forma correta de indexação).

Destacar esses pontos é relevante pois os conceitos de “organicidade” e “tóxico” são criados, alimentados e divulgados, sobretudo, por essas instâncias legitimadoras de segundo grau. Para elas indicar que há links que podem “poluir” o seu desempenho frente às instâncias legitimadoras de primeiro grau (mecanismos de busca) é um dos grandes diferenciais não oferecidos pelas origens de dados como Google Analytics ou Google Search Console. Ele vai ser usado na definição dos discursos veridictórios dos demais sites, como veremos a seguir.

Figura 2. print do website Domain Rating (Ahrefs)



Domain Rating (Ahrefs)

Para a análise do software Ahrefs foi selecionada a página de glossário de suas métricas, fazendo um recorte do texto introdutório e do intertítulo direcionado ao Domain Rating (DR), métrica equivalente ao Authority Score do SEMRush, uma vez que ambos se referem à totalidade do domínio web em questão. Primeiro, gostaríamos de trazer o trecho intitulado “por que você deveria confiar nos dados da Ahrefs”:

Antes de entrarmos na métrica em si, aqui está porque você deveria confiar na Ahrefs. Ahrefs têm o crawler mais ativo da indústria de SEO, de acordo com o CloudflareRadar, e o terceiro crawler mais ativo no mundo, atrás apenas do GoogleBot e do BingBot. Ainda temos infraestruturas top de linha, e somos a única ferramenta usando tecnologias e insights provenientes da construção de nosso próprio mecanismo de busca para melhorar as ferramentas e dados da Ahrefs. (Bek, 2023, p. tradução nossa)

O trecho é sintomático do que viemos analisando. O uso da terceira pessoa para Ahrefs só é quebrado no último período, quando se referem a si mesmos como um “somos” que desenvolvem outras ferramentas, além do Ahrefs, que servem ao propósito de torna-lo mais confiável e fidedigno. No segundo período a autoridade do Ahrefs é baseado em um percurso gerativo de sentido em que o seu crawler – ferramenta que analisa os sites disponíveis online, captando dados provenientes das linguagens HTML e CSS e compilando-os em índices – como o mais ativo.

Percebe-se aqui que há um entendimento dessa comunidade de que um crawler ativo é sinônimo de maior proximidade entre a métrica construída e o índice real e secreto do Google. Quanto mais se aproximam do fazer-secreto dos buscadores – colocados em comparativo como os “únicos mais ativos” – mais validade essa ferramenta ganha por usar as formações secretas de discurso para sua estrutura.

Na construção do discurso veridictório, o enunciado que faz-parecer secreto, ou seja, aquele em que o “ser” e o “parecer” estão em disjunção, é o das camuflagens subjetivantes. Tais camuflagens formulam uma comunidade onde não é necessário

expressar o que é um crawler nem como ele funciona, mas dizer que o seu funcionamento secreto é tão grande quanto o daqueles a quem se quer imitar é suficiente para lhe garantir um lugar no rol de credibilidade que os mesmos mecanismos de busca merecem.

À essa construção discursiva do secreto se sobrepõe uma camada de cientificismo, ou uma segunda camuflagem, dessa vez objetivante, quando se exime o enunciador de dizer isso, atribuindo a outro – no caso, a companhia CloudflareRadar – de ter criado a estatística. Assim, um segredo com indício de ciência veste duas camuflagens e permanece bem encapsulado em um discurso amarrado de significados.

Por fim, a credibilidade se consagra com uma performance desse “nós” ausente até o último período: o momento em que afirmam munir a ferramenta e os dados do que eles encontraram em um processo de “engenharia reversa”, quando se esforçaram para recriar um mecanismo de busca, incorporando aquilo que descobriram em uma ferramenta mais robusta. Não se questiona nem se coloca o sucesso do mecanismo de busca criado – uma tentativa falha poderia munir a ferramenta de resultados inúteis – mas se afirmar que a experiência pela experiência lhes imbui de credibilidade.

O uso de pesquisas aparece mais adiante, quando o “Airas Report: state of link Building in 2022”⁵ é usado como fonte para trazer dois dados, endossando a conjunção entre o sujeito-profissional de SEO e o objeto-métrica da Ahrefs. Primeiro, o número de 64% de profissionais SEO confiarem nos dados de link Building da Ahrefs. Para nossa análise, esse ponto é ímpar. Todo o motivo narrativo do texto é a construção do sentido de credibilidade da métrica que nasce impossibilitada de representar o mundo, dada a natureza do código. Porém, não apenas a pesquisa aponta que essa confiança – ou seja, a performance proposta pela narrativa – já foi alcançada, como também que ela é maior do que a dos dados oficiais (aqueles provenientes das plataformas do Google, o Google Search Console (GSC)).

O segundo valor trazido pela Airas Report é que 44% dos profissionais de SEO trazem as métricas da Ahrefs como suas métricas oficiais. Essa escolha por uma métrica oficial implica uma práxis de um campo, algo que, na narrativa, indica que há uma competência a ser alcançada se o sujeito quiser entrar em conjunção com o campo de profissionais de SEO. Tal conjunção é parte das sanções positivas guardadas aqueles que fazem o percurso gerativo de sentido do texto: que tem contato com a métrica, a entendem, confiam nela e a usam para seus fins profissionais.

As definições da métrica de Domain Rating (DR), citadas entre o URL Rating (outra métrica, essa usada para endereços URL separados) e o Ahrefs Rating (métrica genérica). O DR se apoia na dinâmica do Page Rank, uma métrica criada para orientar o funcionamento do Google, em que a quantidade de links apontando para um website aumentaria a probabilidade de um usuário entediado e de uso errático da web acabar em um site dado o volume de links existentes na web (Brin, Page, 1998). A construção dessa métrica, porém, funcionaria via engenharia reversa: usando o *crawler* potente

⁵ <https://aira.net/state-of-link-building/link-building-techniques-and-tools/>

mentionado acima, a Ahrefs teria acesso a quase o mesmo universo do Google, permitindo-o replicar quase na íntegra o Page Rank e oferece-lo aos profissionais de SEO.

Essa proximidade com a possibilidade de replicar o índice é o fio condutor da narrativa do algoritmo e da companhia, que, no texto que explica o DR, não se exime do uso da primeira pessoa. Afinal, com as bases teóricas já fundamentadas na terceira pessoa no texto anterior, este é o momento em que a narrativa já superou o desafio imposto e começa a conjunção entre o sujeito-profissional de SEO e a sanção que ele alcança: ter sucesso em seu trabalho. Para esse sucesso ele precisa dominar as métricas da Ahrefs, o que só poderá fazer conhecendo-as.

O mesmo procedimento de desmembramento para explicação e construção de uma credibilidade é usado quando afirmam que “há três coisas que precisam ser consideradas no cálculo do DR”. A estrutura dos próximos três parágrafos segue o mesmo modelo: um período negrito no plano de expressão, com uma proposição narrativa que usa a camuflagem objetivante e se aproxima, em estilo e narrativa, do texto científico; e dois períodos em sequência com exemplos práticos relacionados à proposição. Assim o enunciado do texto se faz mais claro, em um nível superior de análise, aproximando-se do enunciatário ao compartilhar do seu universo de referências e necessidades.

Trazemos o seguinte exemplo que ilustra e de onde derivaremos as próximas análises do material:

1. O número de links únicos de referência de um domínio que têm ao menos um link “follow”⁶ apontando para o website. Isso significa que links “nofollow” não ajudarão o DR para o website alvo. Uma vez que os domínios de referência contam como apenas um link por site, links subsequentes de um mesmo website também não ajudarão o DR de um site. (Bek, 2023, p. tradução nossa).

Conjunção criada entre o sujeito-profissional de SEO e o objeto métrica através da credibilidade deste, se direciona o sujeito para os usos que ele pode fazer da métrica em seu dia a dia. Há aqui uma camuflagem subjetivante se construindo, a dinâmica do segredo ou do discurso enunciado para um “iniciado” que se torna familiar com uma nova verdade. As referências do discurso são nichadas, construindo uma credibilidade diferente daquela para atestar a qualidade do dado. Os usos do dado, assim e no

⁶ As referências a links “follow” e “nofollow” são à linhas de código HTML, também chamadas tag. Especificamente esses são atributos possíveis da tag <a>, usadas para criar links entre websites. Os atributos informam a um *crawler* se este deve ou não seguir o caminho indicado pelos links em sua varredura do conteúdo de um website. A presença de um link “follow” fornece credibilidade do site para o mecanismo, enquanto a de um “nofollow” o afasta.

segundo texto, mostram ser objeto da dinâmica do segredo, enquanto a origem dos dois, sujeitos da dinâmica das ciências.

Essa estrutura, tornada clara pela análise, é comum em nossos cotidianos, quando livros como “Storytelling com Dados”, de Knaflitz (2019), e “Como mentir com estatística”, de Heff (2016) figuram entre best-sellers e recomendados por nomes como Bill Gates e o próprio criador do PageRank, Sergey Brin. O uso e a manipulação de dados, aqui usando “manipulação” em seu sentido lato, são manifestações de como essa estrutura de camuflagem subjetivante surge e é usada para proteger a credibilidade de seus usos. Sendo muitas vezes o uso do dado, e não sua criação, a fonte de investimento financeiro – em outras palavras, aquilo pelo que se paga – esses modos de usar, ou táticas (Certeau, 2014) com dados acabam se tornando sujeitos dessa camuflagem subjetivante, uma vez que o segredo pressupõe um percurso gerativo de sentido em que o sujeito entra em conjunção com o secreto para a sanção simbólica: no caso, a de ter seu investimento recuperado.

Figura 3. print do website Domain Authority (MOZ)



Domain Authority (MOZ)

O Domain Authority (DA), métrica criada pela Moz foi uma das primeiras do mercado, lançando em 2004 o primeiro estudo de SEO e o primeiro indicador do tipo. O texto que descreve o DA começa, tal qual os demais, explicando o que o indicador se propõe e o que garante que ele é confiável. Essas informações aparecem em dois parágrafos, respectivamente, que buscam defini-lo e dar-lhe uma origem.

O primeiro afirma que o indicador funciona como uma escala de 1-100 que indica, à medida que o valor se aproxima de 100, qual a probabilidade que ele tem de figurar nas SERP. Ao afirmarem que “é uma métrica desenvolvida pela MOZ” já exercem a camuflagem objetivante para mascarar o enunciador e criar esse sentido de verdade.

No segundo, usa-se a forma passiva do inglês, “is based on” para se referir àquilo que alimenta o indicador. “É baseado em” é uma forma de camuflagem objetivante, em que “é baseado em” camufla o “nós baseamos ele em” – no caso, usa-se o “our Link Explorer web index”. Percebe-se que a presença da terceira pessoa em “nosso” Link

Explorer só é cabível quando um dos sujeitos (DA ou Moz) está como sujeito dado da oração.

Assim, podemos perceber que a dinâmica de quem cria o que faz um jogo de camuflagens, sendo ora objetivante, quando se finge que ou os dados que alimentam a métrica são “dados”, ou o indicador é “dado” frente à qualidade dos dados que são coletados pela Moz. Esse discurso de credibilidade de um dado ou indicador é importante de se ter em mente: quebra-se a estrutura de um dado entre o indicador propriamente dito e o modo como ele é feito, sendo que apenas um deles, por vez, pode receber uma primeira pessoa como sujeito. Nessa troca permite-se que o indicador faça o percurso gerativo de sentido de credibilidade.

Vemos ainda outros atores sendo invocados no discurso, como o “machine learning”, similar a outros indicadores, assim como “algoritmo”, termos que são usados de modo a endossar o distanciamento e o solucionismo de dados para garantir a qualidade e a idoneidade da métrica. Surge então o próximo intertítulo, que se chama “como o DA é calculado” (*how is DA calculated*).

O primeiro parágrafo é onde o contexto social se impõe sobre o fazer discursivo do texto, e onde as relações de poder do campo do SEO determinam manifestações que o enunciador não pode se eximir de colocar. Ao explicar – como o fazem os demais produtores de dados – que “Domain Authority is not a Google ranking factor and has no effect on the SERPs”, em negrito, o enunciador se posiciona como sujeito de um possível processo se afirmasse o contrário. O discurso enunciado, assim, traz marcações das relações de poder e do contexto em que está inserido, fazendo com que mesmo o discurso e a tentativa de validação do seu indicador e de sua credibilização seja afetado por esse agente externo.

Não que isso o impeça de tentar tornar essa situação como uma camuflagem subjetivante. Esse é um procedimento já documentado chamado ato de desmentir: para que um discurso tenha sua credibilidade desmentida e, simultaneamente, se construa a credibilidade do enunciador da nova proposição, o discurso antigo deve ser reordenado de modo a assumir uma camuflagem distinta daquela do novo discurso revestido de credibilidade (Gabry, 2020). Isso acontece logo no segundo parágrafo desse novo intertítulo:

No que tange a atualização do Domain Authority 2.0 do começo de 2019, o cálculo da métrica de DA vem de um algoritmo de machine learning que tenta prever quão frequentemente o Google usa um domínio em seu mecanismo de busca. Se o domínio A é mais provável de aparecer nas SERP do Google do que o domínio B, então devemos esperar o DA do domínio A ser maior do que o do domínio B. Saiba mais sobre a atualização do Domain Authority e como discuti-lo em

apresentação para com seu time, ou explorar em como usar o indicador DA 2.0 neste whitepaper simples⁷. (Moz, 2017)

Aqui, depois de ter ferido a jornada de credibilidade entre o sujeito e seu objeto – a certeza da posição na SERP e a sanção do lucro ao enunciário – o discurso faz o ato de desmentir. Ao expressar que o “domínio A é mais provável de aparecer nas SERP do Google” o sujeito Google aparece expresso e seus domínios e SERP aparecem como frutos do “segredo”. Um segredo que a métrica desvela através dos algoritmos de machine learning e da predição do resultado do mecanismo de busca.

A isso se segue um novo exemplo que busca exemplificar essa natureza comparativa do indicador, seguidos de três intertítulos: “como verificar meu DA”, “qual uma boa média de DA”, “como posso aumentar meu DA” e “por que meu DA mudou”. Em todos esses casos são usadas táticas discursivas em que se mostram exemplos e situações possíveis para o sujeito-profissional de SEO, sobretudo trazendo trechos como “Não há nada como uma pontuação absolutamente ‘boa’, ‘média’ e ‘ruim’ de DA; só há pontuações ‘boas’, ‘médias’ e ‘ruins’ dentro de um contexto”⁸ (Moz, 2017). Tal trecho funciona ao ajudar a conjunção do sujeito da narrativa com sua sanção de riqueza, ao relacionar a falta de resultados ruins. Isso é uma forma de direcionar o sujeito à argumentação do que é um bom resultado de um profissional de SEO para seu cliente.

CONCLUSÕES

Com essas análises foram possíveis extrair algumas conclusões a respeito da forma como os dados e métricas constroem sua credibilidade. As ferramentas como MOZ, SEMRush, Ahrefs, entre outras, desempenham papéis fundamentais ao moldar as estratégias adotadas por profissionais e agências, evidenciando uma dinâmica de poder que transcende a simples manipulação de algoritmos para entrar no terreno da criação e gestão da credibilidade online. A interação entre dados oficiais e os gerados por terceiros cria um ambiente competitivo onde a veracidade e a autoridade dos dados são constantemente postas à prova, sublinhando a importância de uma abordagem crítica e bem-informada por parte dos que atuam neste campo.

As métricas de avaliação como Authority Score, Domain Rating e Domain Authority oferecem uma visão quantificada e aparentemente objetiva da relevância e autoridade de um site, contudo, como demonstrado, estas são construídas sob camadas de estratégias discursivas que visam convencer tanto clientes quanto o mercado de suas validades.

⁷ As of the Domain Authority 2.0 update in early 2019, the calculation of a domain's DA score comes from a machine learning algorithm's predictions about how often Google is using that domain in its search results. If domain A is more likely to appear in a Google SERP than domain B is, then we would expect domain A's DA to be higher than domain B's DA. Learn more about the Domain Authority update and how to discuss it with your team with this presentation, or explore how to use DA 2.0 metrics with this comprehensive whitepaper.

⁸ there is no such thing as an absolutely "good," "average," or "bad" Domain Authority score; there are only scores that are "good," "average," or "bad" within the contexto”

As estratégias greimasianas de verificação, como a camuflagem subjetivante e objetivante, são habilmente utilizadas para conferir um verniz de cientificidade e inevitabilidade aos dados, que são, de fato, produtos de uma elaborada construção social e técnica. Este estudo revelou como as métricas não apenas refletem, mas também moldam o ecossistema do SEO, influenciando decisões e estratégias empresariais de maneira significativa.

Este ato ocorre quando uma métrica ou dado é inicialmente desacreditado para, em seguida, ser revalidado sob uma nova luz, oferecendo uma camada adicional de credibilidade e autoridade. No contexto do SEO, isso se manifesta quando os produtores de métricas, como o Domain Authority, inicialmente negam qualquer impacto direto nas SERPs para depois afirmar sua relevância indireta por meio de correlações e evidências obtidas por aprendizado de máquina e análises preditivas.

Ao empregar esse método, as entidades responsáveis pelas métricas reestruturam o discurso em torno de seus produtos, endossando sua utilidade e precisão. Este ato de desmentir não só serve para reafirmar a autoridade e a validade da métrica em questão, mas também para posicionar os produtores de dados como entidades confiáveis e transparentes no mercado de SEO, que conscientemente esclarecem as limitações e o escopo de aplicação de seus produtos. Este processo reflete uma sofisticada interação entre conhecimento técnico e persuasão retórica, essencial para a sustentação e crescimento da influência dessas ferramentas no campo digital.

O ato de desmentir, quando aplicado no contexto de manipulação e apresentação de dados, assume uma posição central na construção da credibilidade e na gestão de expectativas em diversas esferas de conhecimento e indústria. Este processo transcende a simples correção ou esclarecimento de equívocos, tornando-se uma estratégia discursiva poderosa que visa reforçar a autoridade e a confiança em uma determinada fonte de informação ou entidade. Ao admitir inicialmente uma limitação ou um mal-entendido potencial, e depois apresentar dados ou argumentos que superam essas limitações, os enunciadores reforçam a percepção de transparência e confiabilidade.

Por fim, consideramos que, dentre as táticas usadas pelas narrativas que têm por objetivo afirmar a credibilidade do dado, o uso de agentes externos para validá-las é imprescindível. Nunca o poder de suas máquinas ou a competência de seus códigos é construída pela sua reputação, mas pela credibilidade atribuída a eles por terceiros. Esses terceiros podem ser quaisquer actantes do discurso construídos no enunciado do texto: estudos e outros dados, inteligências artificiais, métodos de machine learning ou outros aparecem como esses auxiliares do discurso no esforço de construir a credibilidade.

Essa análise tem por universo o escopo do mercado de otimização para mecanismos de busca, com suas instituições legitimadoras, produtores culturais e consumidores de papéis cambiantes em uma rede de relacionamentos. É possível perceber como essas relações se manifestam nos percursos gerativos de sentido para a criação de credibilidade de uma métrica – algo que acreditamos possua um grau de generalização

toda vez que um novo “dado” ou métrica são criados. Seja um indicador macroeconômico, a mensuração de aumentos de feminicídio durante governos ou dados relativos ao posicionamento de sites em mecanismos de busca, a construção da credibilidade de um indicador perpassa pela rede na qual aquele dado será utilizado.

AGRADECIMENTOS

Agradecemos à Universidade Federal Fluminense e à Capes pelo amparo no desenvolver da pesquisa, e aos esforços da comunidade de SEO no LinkedIn pela iniciativa transparente de dialogar e trocar estratégias acerca das diferentes métricas usadas.

FINANCIAMENTO

Gostaríamos de agradecer à bolsa Capes de amparo social pela oportunidade de realização e dedicação à pesquisa, esforço sem os quais não teriam sido possíveis as disponibilidades de tempo e cursos para sua execução.

DISPONIBILIDADE DE DADOS

Todos os dados disponibilizados estão disponibilizados nas URLs referentes até a data de envio desse material. Histórico dos sites pode ser encontrado via WayBack Machine: <https://web.archive.org/>.

REFERÊNCIAS

Bek, Rebekah. *Ahrefs' SEO Metrics: A Glossary*. [em linha] 2023. Disponível em: <https://ahrefs.com/blog/ahrefs-seo-metrics/#section10>. Acesso em: 18/04/2024.

Bourdieu, Pierre. *A Economia das Trocas Simbólicas*. 3ª ed. São Paulo: Editora Perspectiva, 1992. 367 p. (Coleção Estudos).

Brin, Sergey; Page, Lawrence. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30, n. 1-7, p. 107-117, 1998.

Broder, Andrei. A taxonomy of web search. *SIGIR Fórum*, 36, n. 2, p. 3-10, 2002.

Campbell, Joseph. *O herói de mil faces*. 1º ed. São Paulo: Pensamento, 2007.

Certeau, Michel de. *A invenção do cotidiano: artes de fazer*. Tradução Alves, E. F. 22 ed. Petrópolis: Vozes, 2014.

Dixon, Stacy Jo. Media usage in an internet minute as of December 2023. p. Tabela, 2023. Disponível em: <https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/>. Acesso em: 11/03/2024.

Escandell-Poveda, Raquel; Papí-Gálvez, Natalia; Iglesias-García, Mar. Técnicas digitales para el estudio de las competencias y perfiles profesionales. *Scire: representación y organización del conocimiento*, 29, n. 1, p. 31-42, 2023.

Gabry, Victor Martini. Entre camuflagens: Uma análise semiótica de “fake news” e memes. Em: *I Encontro Nacional da ABCiber*, 2020, [em linha]. ABCiber, Disponível em: <https://abciber.org.br/simposios/index.php/virtualabciber/virtual2020/schedConf/presentations>.

Gingras, Richard. Labeling fact-check articles in Google News. Google News Initiative, online, 2016. Disponível em: <https://blog.google/outreach-initiatives/google-news-initiative/labeling-fact-check-articles-google-news/>. Acesso em: 14/04/2024.

Greimas, Algirdas Julien. *Sobre o Sentido II: ensaios semióticos*. 1ª ed. São Paulo: Nankin: Edusp, 2014.

Heff, Darrell. *Como mentir com estatística*. Rio de Janeiro: Intrínseca, 2016.

Knafllic, Cole Nussbaumer. *Storytelling com dados: um guia sobre visualização de dados para profissionais de negócios*. Rio de Janeiro: Alta Books, 2019. 256 p.

Kritzinger, Wouter Thomas. *Development of a search engine marketing model using the application of a dual strategy*. Orientador: Weideman, M. 2017. 147 f. (Doutorado) - Faculty of Informatics and Design, Cape Peninsula University of Technology, Cape Town. Disponível em: <https://etd.cput.ac.za/bitstream/20.500.11838/2603/1/197041329-Kritzinger-Wouter%20Thomas-D.Tech-Informatics-FID-2017.pdf>.

Morozov, Evgeny. *Big Tech: a ascensão dos dados e a morte da política*. São Paulo: Ubu Editora, 2018. (Coleção Exit).

Moz. Domain Authority. [em linha] 2017. Disponível em: <https://moz.com/learn/seo/domain-authority>. Acesso em: 22/04/2024.

Nuti, Sudhakar V.; Wayda, Brian; Ranasinghe, Isuru; Wang, Sisi et al. The Use of Google Trends in Health Care Research: A Systematic Review. *PLoS ONE*, 9, n. 10, p. e109583, 2014.

Paul, Prantosh; Aithal, P. S., 2018, *Business Information Sciences With Special Reference to the Digital Marketing and SEO as a Field of Study*. 37-47. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3380510.

Rocha, João Cezar de Castro. *Guerra cultural e retórica do ódio: crônicas de um Brasil pós-político*. Goiânia: Editora e Livraria Caminhos, 2021.

Sahin, Alperen; Dirsehan, Taskin. An Overview of Search Engine Marketing: A Systematic Literature Review. *Galatasaray University Managerial and Social Sciences Letters*, 1, n. 2, p. 125-142, 2023.

Schultheiß, Sebastian, 2023, *How search engine marketing influences user knowledge gain*. ACM. 475-478. Disponível em: <https://dx.doi.org/10.1145/3576840.3578297>.

Sen, Ravi. Optimal Search Engine Marketing Strategy. *International Journal of Electronic Commerce*, 10, n. 1, p. 9-25, 2005/10/01 2005.

Team, Semrush. Semrush Authority Score Explained. [em linha] 2020. Disponível em: <https://www.semrush.com/blog/semrush-authority-score-explained/>. Acesso em: 17/04/2024.

Wang, Caili; Hao, Xiaotong, 2011, Comprehensive Model for Evaluating Search Engine Marketing Effect. In: *Fourth International Conference on Business Intelligence and Financial Engineering*, 2011. 231-234.

Yeoung, Man; Nisar, Tahir. Branding and search engine marketing. In: *5th EMAC Regional Conference: Marketing Theory Challenges in Emerging Markets*, Poland, 2014. [em linha]. Disponível em: <https://eprints.soton.ac.uk/380507/1/EMAC%2520Conference%2520Paper.pdf>.