

CONSTRUÇÃO DO CONHECIMENTO EM TEMPOS ATUAIS: OS PROBLEMAS DA UTILIZAÇÃO DO GOOGLE COMO PRINCIPAL FONTE DE PESQUISA

CONSTRUCTING KNOWLEDGE TODAY: THE ISSUES OF USING GOOGLE AS MAIN SOURCE FOR RESEARCH

André Luiz França Batista*, Bruno dos Santos Simões, Taíse Ceolin
PPGECT - Universidade Federal de Santa Catarina – Florianópolis - SC

Resumo: Este artigo sinaliza algumas das contribuições da Internet para a gestão e construção do conhecimento nos últimos 20 anos. Posteriormente é discutido como empresas privadas, especificamente o Google, tem determinado o modo pelo qual as pessoas tem acesso à vasta quantidade de informação disponível *online*. A invasão de privacidade e o fornecimento de resultados personalizados conforme o perfil de cada pessoa, também são tratados neste trabalho, bem como os riscos decorrentes disto. Por fim, são propostas algumas medidas para que os indivíduos não permaneçam presos a um ciclo de informações repetidas, e assim tenham acesso a conteúdos que antes não estariam ao seu alcance virtual.

Palavras-chave: Construção do conhecimento. Internet. Google. Indutivismo.

Abstract: This paper presents the Internet's contribution to the management and constructing knowledge in the last 20 years. Subsequently, we discuss how private companies, specifically Google, has determined the way people have access to the vast amount of information available online. The invasion of privacy and providing personalized results according to the profile of each person are also dealt in this work as well the risks of this. Finally, we propose some methods to ensure that individuals do not remain trapped in a cycle of repeated information and thus have access to content that would not be before your virtual reach.

Keywords: Constructing knowledge. Internet. Google. Indutivism.

1. Introdução

O acesso a conteúdos informacionais, visando a construção de conhecimentos, tem sido facilitado pelo uso das Tecnologias Digitais de Informação e Comunicação (TDIC) e seus avanços recentes. Desde a sua criação, a Internet tem assumido um papel cada vez mais essencial na gestão, no armazenamento, na distribuição, e no acesso às informações de temáticas diversas. A comunicação entre indivíduos também se beneficiou dos advenços da Internet, tornando o contato entre duas ou mais pessoas mais rápido e eficiente com o passar dos anos (CARR, 2011).

Há hoje um grande volume de informações disponíveis a apenas um clique de distância. Isso ocorre graças às ferramentas de TDIC, chamadas “máquinas de busca”, do inglês *Search Engines* (CROFT, 2010). Tais ferramentas possibilitam a transformação de um dado bruto em conhecimento, em uma fração de segundos. As máquinas de busca, principalmente o Google,

* andreluiz@iftm.edu.br

transformaram-se na principal interface entre a realidade e o usuário/pesquisador na Internet. Contudo, esse fato é alvo de constantes críticas devido aos impactos negativos que decorrem desta intermediação eletrônica.

A complexidade referente aos princípios de democratização do acesso a informação, a dissipação da análise crítica de conteúdos científicos, e a invasão de privacidade das pessoas, são alguns dos problemas motivadores destas críticas a este modelo monopolista de mediação entre o indivíduo e a informação (MAURER, 2007; ZITTRAIN, 2008; SULLIVAN, 2009; CARR, 2011; PARISER, 2012; NHACUONGUE, 2013). De acordo com Sullivan (2009) e Pariser (2012), desde o ano de 2009 o Google coleta informações pessoais dos cidadãos e com base nesses dados, os resultados das buscas são processados, filtrados e entregues ao consumidor em uma lista hierarquicamente organizada.

Pode-se inferir disto que os sujeitos pesquisadores são “induzidos” a absorver as informações entregues na primeira lista de resultados, e tomá-las como fontes singulares na construção do conhecimento, desconsiderando as outras informações que não foram catalogadas pelo Google, ou que foram filtradas e omitidas por não serem compatíveis com o seu perfil. Esse tipo de atitude determinista pode tornar a busca por conhecimentos ainda mais tendenciosa.

Partindo disso, apresentamos aqui, algumas propostas para auxiliar as pessoas no processo de busca de informações e conhecimento, apartando-se de caminhos já definidos por uma máquina ou por uma ferramenta, que as envolve por onde quer que naveguem na Internet.

2. Uma imensidão de informações

Desde o início da década de 1990 a Internet tem causado impacto enorme em nossa cultura e comércio, conforme aponta Schons (2007). O crescimento da comunicação instantânea também foi fomentado pelo uso da Internet com a utilização do correio eletrônico (*e-mails*), mensageiros instantâneos (*instant messengers*), videoconferências, e a *World Wide Web* com seus fóruns de discussão, *blogs*, redes sociais, e *websites* de comércio eletrônico. Warf (2013) relata que tudo isso contribuiu para o aumento do volume de informações que são produzidas, consumidas, e transmitidas a altas velocidades por meio de redes de comunicação.

Conforme estudos estatísticos apresentados no segundo capítulo da obra de Levene (2011), a quantidade de informação disponível na Internet tem uma taxa de crescimento exponencial. Em 2010, estimativas apresentadas por Levene (2011) apontavam que havia 200 milhões *websites* e uma taxa de crescimento de 200 mil de novos *websites* por dia. Em 2015, dados apontam a existência de 1.2 bilhão de *websites* e uma taxa de crescimento de 101 por cento no ano de 2012, e uma queda de 3% no ano de 2013, segundo o *Real Time Statistic Project (2015)*. Esses dados são controversos pois outro *website*, o *World Wide Web Size (2015)*, contabilizava 4.52 bilhões de páginas indexadas em 25 de Fevereiro de 2015. Independente das discrepâncias encontradas acerca da quantidade de *websites* na Internet, podemos concluir que, mesmo assim, a taxa de crescimento é muito alta e que o volume de informações é muito vasto. Contudo, quais são os meios para encontrar e ter acesso à essas informações disponíveis na Internet?

De acordo com Levene (2011), existem diferentes estratégias para encontrar informações na Internet. A escolha de uma ou outra, depende da informação a ser buscada, do nível de sofisticação da pesquisa, e da intencionalidade da busca.

O primeiro, e mais simples, modo de encontrar uma informação na Internet é o acesso direto, que ocorre quando o usuário digita um endereço eletrônico diretamente em seu navegador *web*. O segundo modo consiste em acessar um portal *web* e obter ali as direções para a navegação. Os portais *web* são *websites* que disponibilizam um ponto de partida para o usuário acessar outras informações na Internet, a exemplo de *websites* como UOL, Terra, iG, Yahoo BR, dentre outros.

A grande quantidade de informações disponíveis na Internet, de acordo com Levene (2011), motivou a concepção de meios para buscar e de recuperar essas informações, de um modo rápido e eficiente. Surgiram então, a partir do ano de 1993, as ferramentas conhecidas como “máquinas de busca” que são capazes de catalogar e indexar as informações existentes nos *websites* disponíveis na Internet.

O terceiro meio utilizado para encontrar informações na Internet são estas máquinas de busca cuja qualidade tem evoluído ao longo dos anos, e por isso muitos usuários tem feito de máquinas de busca como Google, Bing, Yandex (Rússia) e Baidu (China), o ponto de partida na Internet para satisfazer as suas necessidades de pesquisa.

Halavais (2013) define máquinas de busca como sistemas de recuperação de informação capazes de realizar buscas por palavras-chave em um texto distribuído digitalmente. Entretanto, essa definição se tornou ainda mais abrangente para incluir também ambientes de armazenamento de informações multimídia tais como fotos e vídeos. Além de realizar pesquisas em textos, as máquinas de busca também são capazes de recuperar informações em imagens e vídeos disponíveis na Internet. Os resultados desta pesquisa de informações são ordenados de modo que aqueles mais relevantes à pesquisa sejam posicionados no topo da lista.

Com a criação do Google, em 1998, houve uma espécie de revolução nos mecanismos de busca, pois o Google utilizava um algoritmo chamado *PageRank* que aumentava a eficiência e qualidade dos resultados obtidos (BRIN, 1998; PAGE, 1999). Isso possibilitou ao Google atingir a posição de máquina de busca mais utilizada do mundo.

É difícil imaginar nossa vida hoje sem o Google. Baibich (2014) aponta o Google, a Internet, os smartphones, e outras tecnologias digitais, como resultados dos avanços científicos da humanidade, e declara assim:

Imagine-se agora sem seu smartphone, sem os muitos gigabytes (ou até terabytes) no seu disco rígido, sem “streaming” de imagens, sem Google na internet (ou mesmo sem internet!) [...]. Tudo isso é resultado dos avanços que aconteceram como consequência da ciência que fizemos, nós, cientistas de todo o mundo, e continuamos fazendo. Foram descobertas científicas que acabaram chegando às ruas de todo o planeta, entrando na vida dos mais ricos aos mais pobres, dos mais sábios aos menos esclarecidos. (BAIBICH, 2014, p. 131).

O pensamento de Baibich (2014) contém ao menos três perspectivas notáveis sobre os avanços tecnológicos, contudo o que mais nos chama a atenção é o entendimento sobre a

dimensão do alcance destes progressos. Ao direcionar a luz para o foco da nossa discussão entendemos – e percebemos em nosso cotidiano – que o Google hoje faz parte da “vida dos mais ricos aos mais pobres, dos mais sábios aos menos esclarecidos”.

Com o foco sobre a marca Google, Haubrich e Carvalho (2013) realizaram um estudo interessante sobre isso e identificaram que ela se apresenta como prática e com qualidade incomparável. De acordo com os autores essas percepções convergem com o que é pretendido pela marca, em suas ações e seu foco de atuação. Haubrich e Carvalho (2013) ainda destacam que é perceptível a supremacia inquestionável do Google frente à concorrência de seu negócio principal, haja vista que ela atende às expectativas de seus usuários de modo consideravelmente satisfatório. Além disso, os autores salientam a simplicidade para localização de serviços e outras informações relevantes, sendo assim:

Não restam dúvidas, que o sucesso do Google está na facilidade quanto à acessibilidade, possivelmente oriundas da transformação social, já que atualmente exigem-se respostas rápidas, além da multiplicidade de funções acumuladas pelos indivíduos, que amplia sua gama de ações, assumindo diferentes papéis, e de sua formação básica e profissional. (HAUBRICH; CARVALHO, 2013, p. 132).

O sucesso do Google pode ser sim oriundo da transformação social, como dizem os autores, entretanto certamente o Google também é um agente transformador social devido ao forte impacto que ele gerou – e ainda gera – em nossas vidas, transformando-as.

Segundo Sullivan (2013), hoje o Google detém mais de 65% das pesquisas realizadas na Internet, e assim tem emergido ao longo dos anos como um líder indiscutível em buscas *online*. O êxito desta ferramenta é tão vasto que, segundo Chen (2009), as expressões “procure no Google”, ou “*Google it*” (em inglês), se tornaram sinônimos de “realizar uma pesquisa na Internet”. Kulathuramaiyer (2006) diz que esta ferramenta de busca se tornou o principal canal de comunicação entre uma pessoa e o resto do mundo, além de transformar o modo de vida das pessoas de muitas maneiras.

3. Uma revolução na internet

Kulathuramaiyer (2006) alega que a nossa percepção é moldada pelo que vemos e pelo que deixamos de ver, e conecta essa alegação com a fala de Weber (2007) que aponta que “a humanidade está em processo de construção da realidade por meio da googlarização”. No original em inglês: “*Mankind is in the process of constructing reality by googeling*”, Weber (2007), refere-se ao ato de ampla utilização do Google como uma ferramenta global. Esse processo de pleno uso desta ferramenta precisa ser analisado cuidadosamente, tendo em vista a possibilidade do direcionamento de ideias que se constitui em uma limitação na construção do conhecimento.

Em 2007, um relatório feito por pesquisadores da Universidade de Graz (Áustria) causou grande impacto na comunidade de ciberativistas e defensores da liberdade de acesso a informação, ao expor como o comportamento monopolista do Google está ameaçando o modo de ver o mundo, e como os seus usuários são tratados como indivíduos completamente sem privacidade. O relatório de Maurer (2007) mostra que atualmente o Google determina a

maneira como procuramos e encontramos informação na Internet utilizando seu algoritmo de ranqueamento (do inglês *ranking*) ou ordenação de resultados.

Langville (2011) esclarece que o Google dispõe de uma relação de bilhões de páginas classificadas em ordem de importância, ou seja, cada *website* tem sua relevância para a Internet como um todo. Este repertório de páginas é composto desde os *websites* mais importantes até os menos relevantes. Esse método de ordenação das páginas, de acordo com Langville (2011), é democrático e reflete o que a “Internet pensa” a respeito de algum assunto ou termo de busca.

Em dezembro de 2009, o Google anunciou em seu *blog* oficial (OFFICIAL GOOGLE BLOG; 2009), a atualização mais significativa no conjunto de suas ferramentas, que ficou conhecida como “busca personalizada”. Até então, o Google retornava os mesmos resultados de pesquisa para todos os usuários. A partir desta data, o conjunto de resultados oriundos de uma pesquisa diferem conforme o perfil do usuário que realiza a busca, ou seja, o Google retorna os resultados das pesquisas dos usuários de modo individualmente personalizado.

Pariser (2012) afirma que usuários distintos recebem resultados distintos a partir de um termo de busca em comum. A coleção de algoritmos, técnicas e ferramentas do Google, agora sugerem os resultados individualmente para cada usuário em particular, ou seja, cada um visualizará resultados completamente diferentes. Assim, não existe mais padronização nas pesquisas realizadas no Google de modo que os resultados são estabelecidos conforme o perfil que este mecanismo traçou para cada pessoa.

Em outras palavras, um algoritmo ordenador de uma empresa privada com ações na bolsa de valores, dita de forma imperativa quais informações o usuário deve receber e quais serão negligenciadas, ou propositalmente suprimidas. Conforme as palavras de Maurer (2007, p. 20), “esta situação é única na história da humanidade”, e denota a forte influência capitalista e de interesses de grandes corporações no modo como as pessoas buscam informações e conhecimentos na Internet.

4. Moldar a visão da realidade

O Google é a principal ponte de acesso entre o usuário e um conteúdo disponível na Internet (KULATHURAMAIYER, 2006; CARR, 2008; CHEN, 2009; CARR, 2011; LANGVILLE, 2011). Baseado nisso autores como Battelle (2005), Maurer (2007), Zittrain (2008), Pariser (2012), Nhacuongue (2013) afirmam que o Google influencia fortemente na construção do conhecimento na atualidade. Como acontece essa influência? Uma das possibilidades que destacamos é o fato dos algoritmos do Google determinarem quais informações serão mostradas ao usuário, e quais serão omitidas.

Na tentativa de oferecer os melhores resultados a cada usuário individualmente, o Google coleta informações pessoais, às vezes sem o consentimento do mesmo. Com isso, o Google tenta “aprender” qual é o perfil de cada usuário, o que ele “gosta” e o que “não gosta”. Eli Pariser, autor do livro “*The Filter Bubble*” (“O filtro invisível”, título na versão em português) afirma em sua obra que o Google coleta 57 informações pessoais dos usuários antes mesmo de realizar a pesquisa requisitada. Pariser (2012) não apresenta uma lista detalhada sobre quais são essas 57

informações. Entretanto, podemos fazer algumas suposições baseadas nas potencialidades de determinadas ferramentas do Google.

Alguns exemplos de informações pessoais que o Google pode coletar de cada usuário são: histórico dos termos de buscas; localização geográfica (com precisão de mais ou menos 2 metros); data e hora da pesquisa; se houve acesso as páginas subsequentes de resultados; tempo gasto para acessar os resultados de uma pesquisa; quais são os *websites* mais visitados; tempo gasto em cada pesquisa; tempo gasto em cada *website*; tempo total de navegação na Internet; navegador utilizado; sistema operacional utilizado (*Windows, iOS, Android, Linux*, outros); o dispositivo utilizado (*desktop, notebook, smartphone, tablet*, outros); idioma do usuário; idiomas dominados (ou não dominados) pelo usuário; frequência de acesso as propagandas veiculadas pelo próprio Google, etc. Estas são apenas suposições sobre quais dados pessoais o Google pode captar dos seus usuários e utilizar na formulação de perfis individuais, considerando que a aquisição destas informações é computacionalmente trivial.

Segundo Pariser (2012), é fácil dar as pessoas aquilo que elas acreditam que desejam, entretanto ao mesmo tempo não é agradável receber as mesmas informações repetidamente. Em seu livro, Pariser (2012) faz uso da expressão “a bolha de filtros” (*The filter bubble*, em inglês) para explicar o universo pessoal *online* de informações no qual nós vivemos. Segundo o autor, esse universo é construído estritamente para cada usuário por meio da utilização de filtros invisíveis personalizados que se encontram na Internet. Essa bolha de filtros é um lugar confortável por definição, pois é povoado de conteúdos que frequentemente acessamos. No entanto, isto também é um problema real uma vez que o conjunto de conteúdos que gostamos de acessar, ou consumir, não é o mesmo conjunto de informações que necessitamos conhecer para construir um determinado conhecimento.

Ao apresentar sua crítica ao Google, Nhacuongue (2013) foca no sistema de hierarquização dos resultados que, segundo o autor, é baseada em interesses comerciais, pois grande parte das informações priorizadas na lista de resultados referem-se a corporações que veiculam suas ações de *marketing* nessa máquina de busca. Assim, segundo o autor, aqueles usuários que desconhecem totalmente o assunto de pesquisa, ou aqueles inexperientes em realizar buscas na Internet, são induzidos a consumir as informações que o Google seleciona.

O Google omite as opiniões que discordam do perfil do usuário. O conjunto de algoritmos oferece economia de tempo na busca pelos assuntos que nos interessam, porém, como alerta Pariser (2012), esses filtros reduzem o contato com visões divergentes, que seriam ingredientes essenciais para fortalecer o debate democrático e estimular o interesse por novos conhecimentos, para não restringir os caminhos dos usuários ou prendê-los na repetição de temas e no viés de confirmação. O Google e outros buscadores não ponderam o sentido do que buscam, não fazem nenhuma análise do tipo que faria um especialista no assunto buscado, sendo assim é necessário que o usuário tenha uma visão crítica para interpretar a lista de resultados fornecida pelo mecanismo de busca.

A construção do conhecimento não pode ser um ato tão “viesado”, polarizado em atitudes e posições com crenças persistentes. Além disso, esta não pode ser uma jornada que tem seus

caminhos induzidos por uma empresa privada ou por suas técnicas e ferramentas computacionais. O usuário deveria ter autonomia para selecionar as informações que efetivamente lhe interessam para construção de seu conhecimento.

5. O Google como ferramenta de pesquisa

O Google disponibiliza diversas ferramentas para seus usuários, das mais simples às mais avançadas. Mas aquela que se tornou uma das mais usadas do mundo é a ferramenta de busca. Atualmente esta ferramenta é comumente usada como principal fonte de pesquisa pelos estudantes, seja em nível superior (BARBASTEFANO; SOUZA, 2007; NICCHELLATTI et al., 2009) ou da educação básica (FIALHO, 2013). Isso remete a uma necessidade eminente de compreender melhor como essa máquina de busca funciona, e também entender como podemos orientar os estudantes a buscarem informações nestas ferramentas.

Nicchellatti (2009) e seus colaboradores investigaram sobre o grau de acessibilidade à Internet e os tipos de *websites* mais visitados pelos estudantes das séries iniciais do curso de Administração de duas Instituições de Ensino Superior do sul do Brasil. Os autores concluíram que é importante incentivar o uso direcionado de buscas em portais como, por exemplo, o Google Acadêmico, pois esta ferramenta oferece acesso a periódicos, bem como a teses e dissertações.

Fialho (2013) apresenta uma possibilidade para essa orientação aos estudantes sobre a pesquisa na Internet. A autora discute o termo “pesquisa escolar orientada”, presente na obra de Carol Kuhlthau. Entre os objetivos da pesquisa está o de orientar a busca e o conceito de fontes de informação para estudantes de ensino médio, nos suportes tradicionais (livros, manuais, revistas) e digitais. Apesar do foco não ser necessariamente a pesquisa no Google, foi indicado ao longo da investigação que esta era uma fonte de pesquisa muito utilizada pelos estudantes. Fialho conclui que é importante ensinar os estudantes a pesquisarem e a avaliarem criticamente a qualidade das informações obtidas.

Sob outra perspectiva, Barbastefano e Souza (2007, p.2) discutem sobre o uso da Internet no plágio de trabalhos acadêmicos e indicam fatores que são apontados na literatura como possíveis causas para isso: “facilidade de acesso à informação, falta de análise crítica em relação aos textos e dificuldades de parafrasear”. Dentre os vários aspectos apontados destacamos a questão ligada à análise crítica em relação aos textos. Essa é uma decorrência, também, das fontes de pesquisa utilizadas pelos estudantes investigados. Os autores apontam que os *websites* que o Google coloca no topo da lista de resultados são os mais acessados pelos estudantes.

Em outra direção, Passerino (2010, p.68) faz uma reflexão sobre a função social das tecnologias no processo educativo. Neste trabalho, o autor faz uma crítica a concepção de que a tecnologia é detentora do conhecimento, e que este pode ser transmitido ao aprendiz como algo externo ao sujeito e acabado, que pode ser “embutido” dentro da tecnologia e desta forma “aprendido” pelo aluno. Para tanto, Passerino (2010) argumenta que as tecnologias deveriam ser usadas como mediadores do conhecimento, mas que é preciso ter clara a não neutralidade dessas ferramentas. Desta forma, entendemos que o uso intencional e direcionado é crucial na

promoção da autonomia do sujeito em relação às tecnologias e na apropriação de conhecimento.

A análise que queremos indicar aqui diz respeito à análise crítica que deve ser feita nas fontes de pesquisa disponibilizadas por meio do Google. Tendo em vista que os resultados apresentados são altamente tendenciosos, uma pesquisa, especialmente acadêmica, não pode contemplar aspectos tão partidários. E como a construção de um trabalho acadêmico demanda na leitura crítica de materiais, a forma de buscar estas fontes não pode ser diferente. Ressaltamos a importância da utilização de bases de dados como a da CAPES <<http://www.periodicos.capes.gov.br>>, SciELO <<http://www.scielo.org>>, ou mesmo o Google Acadêmico <<http://scholar.google.com>>. No entanto, devemos lembrar que os estudantes, inclusive do ensino superior, ainda utilizam o Google como principal fonte de pesquisa sem a análise crítica, e mesmo sem saber da “manipulação” personalizada dos resultados que obtém.

Além da compreensão de que os resultados obtidos em uma busca carregam consigo toda uma bagagem de intenções, em grande parte comerciais, é importante destacarmos o processo de indução pelo qual passamos ao realizarmos consultas em ferramentas *online* como o Google.

6. Pesquisas personalizadas: resultados indutivistas

As ferramentas disponíveis na Internet evoluem constantemente e tornam-se mais eficientes ao longo do tempo, adaptando-se às necessidades dos usuários. Entretanto, tais melhorias precisam ser analisadas com a devida atenção e cuidado, tendo em vista que por mais eficientes que sejam, estão de acordo com as intencionalidades, nem sempre explícitas, de uma empresa privada. Outro ponto a ser considerado é a possível alienação ou desatenção do indivíduo que utiliza tais ferramentas, especificamente do Google, sem dar-se conta da indução “disfarçada” dos mecanismos de busca que disponibilizam as informações que nem sempre correspondem ao que o usuário necessitaria.

Quando o Google tenta traçar o perfil de cada usuário baseado em seus comportamentos virtuais e características pessoais, essa representação pode não condizer com a realidade. Isso expressa um problema elementar da indução, que é o mecanismo fundamental utilizado pelos algoritmos para realizar previsões.

Filósofos já refletiam sobre esse tema bem antes de surgirem computadores capazes de realizar induções. Mesmo que seja possível demonstrar a verdade de uma prova matemática por meio dos princípios iniciais da indução, o filósofo David Hume (1972) enfatiza que a realidade não se comporta desse modo. Basicamente, desempenhos anteriores não prognosticam boas premissas para o futuro.

Em outras palavras, o problema da indução refere-se à generalização sobre as características de um tipo de objeto de acordo com observações sobre um determinado número de exemplares desse tipo de objeto. Ou ainda refere-se a inferência de que uma sequência de eventos futuros acontecerá conforme foi observado anteriormente.

Hume (1972) ressalta ainda que:

Em verdade, todos os argumentos derivados da experiência se fundam na semelhança que constatamos entre objetos naturais e que nos induz a esperar efeitos semelhantes àqueles que temos visto resultar de tais objetos. Apesar de somente um bobo ou um louco – e ninguém mais! – pretender discutir a autoridade da experiência ou rejeitar este grande guia da vida humana [...] (HUME, 1972, p.39-40).

Outro pensador que dedicou-se a investigar e tentar resolver o problema da indução foi Karl Popper (1902-1994). Um dos focos de seu trabalho estava na premissa de que há teorias e ideias falhas que se encaixavam perfeitamente com o método científico, mas que estavam equivocadas.

Popper (2008) colocou o problema sob uma outra perspectiva: “se somente encontramos cisnes brancos ao longo do tempo não implica que todos os cisnes sejam brancos”. De acordo com ele, é necessário procurar por um cisne negro, ou seja, por um contraexemplo que falseie a conjectura. A falseabilidade é fundamental na busca pela verdade.

A finalidade da ciência, para Popper (2004), é elaborar afirmações para as quais não se encontre exemplos contrários. A perspectiva de Popper inspira certa humildade acerca do conhecimento induzido cientificamente, ou seja, é preciso compreender que podemos estar equivocados com tanta frequência quanto estamos corretos, e usualmente não percebemos quando isso ocorre.

A ausência dessa "humildade" é o que transforma muitos métodos algorítmicos de previsão em ferramentas indutivistas. Essa tendência atual pode reforçar mentes partidárias e limitadas, bem como colocar o usuário em meio a um “vale sombrio” que ecoa informações repetidas. Contudo, podemos apontar caminhos para uma diversidade de perspectivas distintas afim de contornar essa tendência e sair desse "entremontes"¹.

7. Fuga do vale da sombra da indução: algumas considerações

Até esse ponto da discussão argumentamos que o crescimento dos filtros incorporados às pesquisas *online* está modificando o modo como buscamos conhecimento na Internet, e em uma instância mais recente, como compreendemos o mundo. Muitos estudantes e pesquisadores não tem consciência de como a Internet funciona, e desconhecem que as empresas de busca – especialmente o Google – determinam quais informações serão vistas e quais serão omitidas, baseado em um perfil traçado com informações pessoais de cada um.

É de essencial importância realizar ações para conscientização de estudantes, pesquisadores, educadores, sobre a imensidão de informações que estão disponíveis na Internet e que não chegam até nós. De certa forma realizar divulgação científica utilizando redes sociais

¹ Vale: Lugar rodeado por montes ou montanhas. Nesse sentido, significa ampliar as visões e perspectivas do usuário para além daquilo que lhe foi induzido pelo mecanismo de busca.

como Facebook, e meios de divulgação de vídeos como o Youtube, que podem ser importantes aliados na implementação deste tipo de ações. A elaboração e disseminação de materiais informativos sobre os filtros personalizados, pode contribuir para formação de uma rede de indivíduos mais cômicos dos riscos deste cerceamento de informações.

Partindo deste pressuposto da conscientização, o próximo passo seria mostrar que existem outras opções, outros meios para alcançar as informações desejadas. São outros caminhos que podem levar a outros destinos. A adoção de uma postura não acomodada, não conformada, e não cativa a apenas uma máquina de busca, pode abrir uma janela de acesso à conteúdos informacionais interessantes que antes não estavam no radar das pessoas. Os resultados das buscas precisam ser contemplados sob um olhar crítico, independentes de qual foi a máquina de busca utilizada.

Uma outra proposta para escapar dos "entremontes" *online* refere-se à previsibilidade do comportamento do pesquisador. Desde o início dos anos 2000, pesquisas já eram realizadas acerca do comportamento dos usuários das máquinas de busca, conforme mostra o trabalho de Hölscher (2000). Abster-se de condutas previsíveis pode contribuir com o código de personalização dando-lhe mais amplitude para cumprir o seu papel. De acordo com autores como Goel (2010) e Bateman (2012) as pessoas assumem atitudes previsíveis ao buscar informações, como por exemplo: acessar os mesmos *websites*, utilizar invariavelmente a mesma máquina de busca, ler frequentemente os mesmos jornais *online*, ouvir as mesmas rádios *online*, interagir virtualmente com as mesmas pessoas, etc.

Não é simples fugir destes ciclos de repetição. Não é fácil executar uma mudança de hábitos. Porém, assim como escolher um novo caminho para chegar até o local de trabalho, nos faz conhecer melhor a cidade em que vivemos, variar nossos caminhos *online* aumenta a chance de nos deparar com novas ideias e novas pessoas. A alteração da rotina de uso da Internet é uma atitude pessoal que dificultará a rotulação do indivíduo. Divergir o foco, mudar direções, colaborará para que o sujeito amplie a sua compreensão de mundo. Essa mudança de hábitos pode parecer assustadora no início mas o contato com novas ideias, pessoas e culturas é cativante.

O argumento de Pariser (2012) é que a indução algorítmica que nos cerca pode levar a um determinismo de informações, no qual o nosso passado determinará o nosso futuro. Segundo o autor, se as pessoas não modificarem suas rotinas, e conseqüentemente os seus históricos virtuais, elas estarão fadadas a viver em repetição.

Uma fração do plano de fuga para escapar do cerceamento provocado pelos filtros, reside na mudança de comportamento das pessoas em suas rotinas *online*. Porém isso terá conseqüências limitadas se as empresas que promovem a personalização não mudarem o seu comportamento também.

Na base de toda problemática está o fato de que pela primeira vez, uma ferramenta tecnológica é capaz de descobrir quem somos, aprender sobre o que gostamos e deduzir o que queremos. Ainda que esse código de personalização não seja infalível, ele é preciso o suficiente para ser lucrativo. Embora a Internet nos forneça acesso a uma vasta quantidade de fontes e

opções, quando temos nosso perfil inferido por um modelo computacional, nós perdemos muitas delas devido ao sobreajuste¹ do modelo.

Um meio para evitar esse sobreajuste, como sugere Popper (2008), é tentar provar que o modelo está incorreto construindo algoritmos que “contestem” o perfil *online* das pessoas visando ajustá-lo. Vejamos uma “proposta-exemplo” a seguir.

Considere este cenário: um cidadão brasileiro, residente no Brasil, cuja língua nativa é o português, realiza uma busca qualquer no Google. Se ele receber algum resultado em francês e passar a acessar frequentemente *websites* em língua francesa, o Google mostrará outros resultados similares e “entenderá” (hipótese) que esse sujeito domina os idiomas português e francês. Porém se o Google desejar ter uma imagem que melhor represente este cliente, deveria constantemente testar essa conjectura sugerindo *websites* em outras línguas, na tentativa de provar que a hipótese estava equivocada. Caso contrário, o cidadão pode se encontrar preso em um vale povoado apenas por *websites* brasileiros e franceses.

Sendo assim, Pariser (2012) apresenta como uma possível sugestão a implementação de critérios de “falseabilidade” nos algoritmos de busca, para testar e refutar os perfis traçados com base nas características das pessoas. Essa prática abrirá uma porta de acesso para serviços, *websites* e conteúdos informacionais que antes estava fechada, mas que mesmo assim pode ser interessante devido a curiosidade pelo inexplorado.

8. Considerações finais

As propostas aqui apresentadas, tanto aquelas centradas no indivíduo quanto aquelas com foco nas empresas, visam propiciar às pessoas o acesso a temas distantes de experiências usuais para aprimorar a construção do conhecimento. Além disso, são sugestões eficientes que mantêm o indivíduo focado no processo de constituição de saberes, e ao mesmo tempo lhe apresenta novos temas, ideias, conceitos e pessoas. São estes novos e diferentes conteúdos que nos inspiram e fazem da construção do conhecimento uma jornada mais interessante e encantadora.

9. Referências

BAIBICH, M. N. **Uma descoberta que mudou a forma de pensar, ou como a ciência muda a humanidade.** *Ciência e Natura*. Edição Especial 35 anos – Vol. I, p. 130-136. 2014.

BARBASTEFANO, R. G.; SOUZA, C. G. **Plágio em trabalhos acadêmicos: uma pesquisa com alunos de graduação.** In: *XXVII Encontro Nacional de Engenharia de Produção*, Foz do Iguaçu, 2007.

¹Do inglês *overfitting*, é o termo utilizado quando um modelo matemático ou computacional se ajusta demais ao conjunto de dados e amostras, assim este não é uma boa representação da realidade.

Disponível em: <http://www.abepro.org.br/biblioteca/enegep2007_tr660482_9513.pdf>. Acesso em 21 Jan. de 2015.

BATEMAN, S.; TEEVAN, J.; WHITE, R. W. **The search dashboard: how reflection and comparison impact search behavior.** In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. p. 1785-1794, 2012.

BATTELLE, J. **The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture.** New York: *Portfolio*, 2005.

BRIN, S.; PAGE, L. **The anatomy of a large-scale hypertextual Web search engine.** *Computer Networks and ISDN Systems*, v. 30, n. 1, p. 107-117, 1998.

CARR, N. **A geração superficial. O que a internet está fazendo com os nossos cérebros.** Rio de Janeiro: Agir, 2011.

_____. **A grande mudança: Reconnectando o mundo, de Thomas Edison ao Google.** Landscape, 2008.

CHEN, R.; KRAEMER, K. L.; SHARMA, P. **Google: The World's First Information Utility?** *Business & Information Systems Engineering*, v. 1, n. 1, p. 53-61, 2009.

CROFT, W. B.; METZLER, D.; STROHMAN, T. **Search engines: Information retrieval in practice.** Reading: Addison-Wesley, 2010.

FIALHO, J. **Experiência com estudantes do ensino médio através da pesquisa escolar orientada.** *Perspectivas em Ciência da Informação*. v.18, n.1, p.15-25, jan./mar. 2013.

GOEL, S. **Predicting consumer behavior with Web search.** *Proceedings of the National Academy of Sciences*, v. 107, n. 41, p. 17486-17490, 2010.

HALAVAIS, A. **Search engine society.** John Wiley & Sons, 2013.

HAUBRICH, G. F.; CARVALHO, C. S. **Google na cabeça! Um estudo sobre o significado da marca.** *Cadernos de Comunicação*, v. 16, n. 2, p. 119-139. 2013.

HÖLSCHER, C.; STRUBE, G. **Web search behavior of Internet experts and newbies.** *Computer networks*, v. 33, n. 1, p. 337-346, 2000.

HUME, D. **Investigação acerca do entendimento humano.** Coleção Biblioteca Universitária Série 1º - Filosofia. São Paulo: Nacional, 1972.

KULATHURAMAIYER, N.; BALKE, W. T. **Restricting the View and Connecting the Dots – Dangers of a Web Search Engine Monopoly.** *J. UCS*, v. 12, n. 12, p. 1731-1740, 2006.

LANGVILLE, A. N.; MEYER, C. D. **Google's PageRank and beyond: The Science of search engine rankings.** Princeton University Press, 2011.

LEVENE, M. **An introduction to search engines and web navigation.** John Wiley & Sons, 2011.

MAURER, H. **Report on dangers and opportunities posed by large search engines, particularly Google.** Graz University of Technology. 2007.

NHACUONGUE, J. A.; FERNEDA, E. **A construção do conhecimento na atualidade: um olhar sobre o impacto da tecnologia.** *PBCIB*, v. 8, n. 1, 2013. Disponível em: <http://www.dgz.org.br/dez12/Art_05.htm>. Acesso em: 15 Jan. 2015.

NICHELLATTI, T. P.; ZAGUINI, J. A.; DOMINGUES, M. J. C. de S.; MAINARDES, E. **Análise do grau de acessibilidade e forma de utilização da tecnologia da informação na perspectiva de estudantes universitários.** In: *VI CONVIBRA – Congresso Virtual Brasileiro de Administração*, 2009. Disponível em: <http://www.convibra.com.br/2009/artigos/199_0.pdf>. Acesso em 30 Jan. 2015.

Official Google Blog. **Personalized Search for everyone.** Official Google Blog, 2009. Disponível em: <<http://googleblog.blogspot.com.br/2009/12/personalized-search-for-everyone.html>>. Acesso em: 22 Jan. 2015.

PAGE, L. **The PageRank citation ranking: Bringing order to the web.** Relatório técnico. *Stanford InfoLab*, 1999. Disponível em: <<http://ilpubs.stanford.edu:8090/422/>>. Acesso em: 22 Jan. 2015.

PARISER, E. **O filtro invisível: o que a internet está escondendo de você.** Zahar, 2012.

PASSERINO, L. M. **Apontamentos para uma reflexão sobre a função social das tecnologias no processo educativo.** *Texto Digital*, v. 6, n. 1, 2010.

POPPER, K. R. **A lógica da pesquisa científica.** Editora Cultrix, 2004.

_____. **Conjecturas e refutações: o progresso do conhecimento científico.** Ed. Universidade de Brasília, 2008.

REAL TIME STATISTIC PROJECT. **Internet Live Stats.** 2015. Disponível em: <<http://www.internetlivestats.com/total-number-of-websites>>. Acesso em: 21 Fev. 2015.

SCHONS, C. H. **O volume de informações na Internet e sua desorganização: reflexões e perspectivas.** *Informação & Informação*, v. 12, n. 1, 2007.

SULLIVAN, D. **Google Still World's Most Popular Search Engine By Far, But Share Of Unique Searchers Dips Slightly.** *Search Engine Land*, 2013. Disponível em: <<http://searchengineland.com/google-worlds-most-popular-search-engine-148089>>. Acesso em: 16 Jan. 2015.

SULLIVAN, D. **Google's Personalized Results: The "New Normal" That Deserves Extraordinary Attention.** *Search Engine Land*, 2009. Disponível em: <<http://searchengineland.com/googles-personalized-results-the-new-normal-31290>>. Acesso em: 22 Jan. 2015.

WARF, B. **Origins, Growth, and Geographies of the Global Internet.** In: *Global Geographies of the Internet*. Springer Holanda. p. 9-44, 2013.

WEBER, S. **Das Google-Copy-Paste-Syndrom.** *Wie Netzplagiate Ausbildung und Wissen gefahrden.* Telepolis, 2007.

WORLD WIDE WEB SIZE. **The size of the World Wide Web: The Internet.** 2015. Disponível em: <<http://www.worldwidewebsite.com>>. Acesso em: 21 Fev. 2015.

ZITTRAIN, J. **The future of the internet – and how to stop it.** Yale University Press, 2008.