

Uma leitura pragmaticista e comunicacional sobre a exploração de dados e mediações algorítmicas – o caso Cambridge Analytica

Marcelo Hamdan Alvim | marcelo.h@me.com

Doutorando pelo PPGCOM/ECA-USP e mestre em Comunicação Social pela PUC Minas

resumo

O presente artigo busca apresentar uma leitura, sob o ponto de vista pragmaticista de C. S. Peirce, das ferramentas algorítmicas de *big data* contextualizadas pelo caso da Cambridge Analytica e sua possível interferência nas eleições americanas de 2016 e no referendo conhecido como Brexit, no mesmo ano. A dinâmica de funcionamento dessas ferramentas é analisada pelos conceitos de Objeto Dinâmico, Objeto Imediato, suas correlações com a formação de crenças e hábitos e seu papel na previsão e na causação de determinados cursos de ação.

Palavras-chave: Big data. Cambridge Analytica. Objeto. Hábito. Semiótica. Peirce.

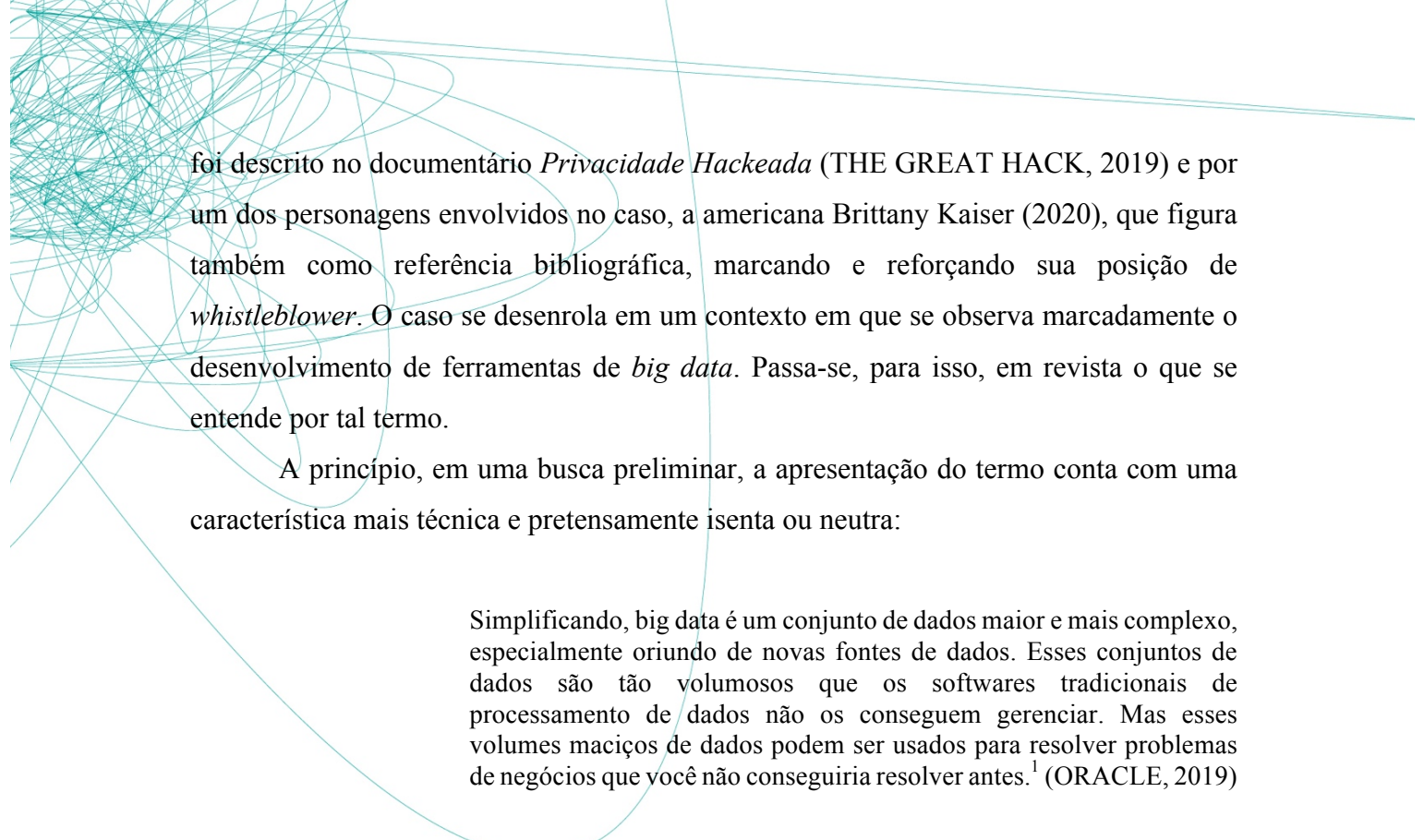
abstract

This article presents a point of view, under the pragmatic theory of C. S. Peirce, of the big data algorithmic tools contextualized by the case of Cambridge Analytica and its possible interference with the 2016 American elections and the referendum known as Brexit, in the same year. The dynamics of how such tools work is analyzed under the concepts of Dynamic Object, Immediate Object, their correlations with the formation of beliefs and habits and their role in predicting and causing certain courses of action.

Keywords: *Big data. Cambridge Analytica. Object. Habit. Semiotics. Peirce.*

Introdução

As complexas circunstâncias envolvendo eventos como a eleição de Donald Trump, em 2016, e a realização do referendo conhecido como Brexit, no mesmo ano, trazem, em meio a diversos fatores e atores, a figura da empresa Cambridge Analytica. A companhia, como se soube após ampla e escandalosa repercussão na mídia, operou – com obtenção e análise de dados de usuários, principalmente da rede social *online* Facebook – a construção de perfis com informações de cunho pessoal, psicológico e comportamental. O caso, além de divulgado e debatido nas esferas de notícias, também



foi descrito no documentário *Privacidade Hackeada* (THE GREAT HACK, 2019) e por um dos personagens envolvidos no caso, a americana Brittany Kaiser (2020), que figura também como referência bibliográfica, marcando e reforçando sua posição de *whistleblower*. O caso se desenrola em um contexto em que se observa marcadamente o desenvolvimento de ferramentas de *big data*. Passa-se, para isso, em revista o que se entende por tal termo.

A princípio, em uma busca preliminar, a apresentação do termo conta com uma característica mais técnica e pretensamente isenta ou neutra:

Simplificando, big data é um conjunto de dados maior e mais complexo, especialmente oriundo de novas fontes de dados. Esses conjuntos de dados são tão volumosos que os softwares tradicionais de processamento de dados não os conseguem gerenciar. Mas esses volumes maciços de dados podem ser usados para resolver problemas de negócios que você não conseguiria resolver antes.¹ (ORACLE, 2019)

Inclusive, na esteira desse tipo de conceito e diante da magnitude da disponibilidade de dados, há proposições a respeito de uma espécie de dispensa de estudos teóricos sobre temas para os quais podem vir a ser obtidos dados, uma vez que há a pretensão de se ter uma quase total detecção de elementos para se descrever determinado objeto, em uma discussão acerca do “fim da teoria” (ANDERSON, 2008; KITCHIN, 2014).

Por ser objeto de estudo transversal, o *big data* é extensamente discutido e suas características principais são exaustividade (totalidade do objeto sendo capturada, ao invés de obtida por amostragem), refinamento (alta resolução ou alta característica indicial), relacionalidade (tem campos comuns com outras bases de dados), extensividade (seus campos podem ser alterados ou aumentados facilmente) e escalabilidade (pode crescer rapidamente), veracidade (pela sua quantidade de incerteza e erro), valor (pela potência em gerar *insights*) e variabilidade (o sentido dos dados pode variar de acordo com a evolução do contexto) (KITCHIN e MCARDLE, 2016). Mesmo assim, não é a simples captação e disponibilidade dessa informação que configura o *big data*, mas “[...] é só quando milhões e bilhões de parcelas individuais de dados são postas juntas

¹ Tradução nossa do original em inglês: “Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can’t manage them. But these massive volumes of data can be used to address business problems you wouldn’t have been able to tackle before.”

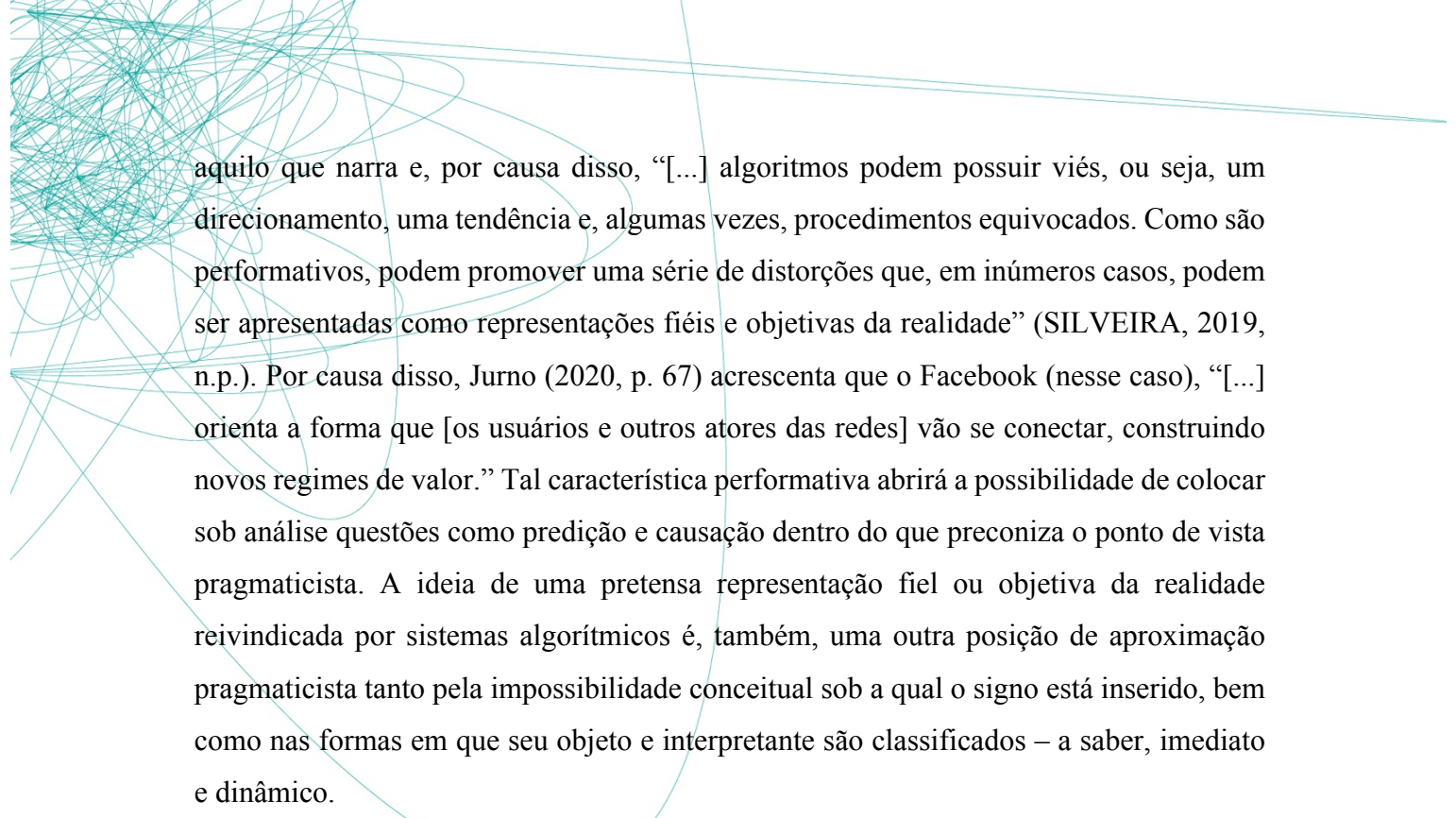
algoritmicamente que o *commodity* conhecido como *big data* emerge.”² (THATCHER; O’SULLIVAN; MAHMOUDI, 2016).

Entretanto, o que se pode apontar é a relevância da consideração de um ponto de vista que suspenda a pretensão de isenção ideológica ou designação meramente comercial desse tipo de volume de informações. De fato, a obtenção da quantidade de dados para a constituição dessa massa chamada *big data*, bem como sua exploração ou tratamento, passam pela construção de sub-base de dados ou identificação de padrões operados por rotinas lógicas e códigos algorítmicos. Algoritmos são códigos (até nova ordem, pois isso parece estar em disputa por conta do desenvolvimento e aprimoramento das inteligências artificiais) concebidos com intencionalidades específicas, escritos por profissionais ou corporações com agendas específicas. Por conta dessa constatação inicial, já parece ser possível dispensar a neutralidade desse objeto (THATCHER; O’SULLIVAN; MAHMOUDI, 2016).

A construção da discussão que se apresenta passa pelas estratégias de conformar ou modular ações dos usuários de redes sociais *online* por parte da Cambridge Analytica, considerando que “[a] captura dos dados se torna a captura das subjetividades”. A possível apropriação de características de certos usuários das redes sociais *online* por lógicas algorítmicas, que mineram e processam os dados capturados que formam parte da massa entendida como *big data*, forma o contexto da presente discussão. A natureza do algoritmo deve ser, portanto, levada em conta, uma vez que são ferramenta-chave na operação do processo que se apresenta. Tais códigos computacionais são constituídos de “[...] instruções inequívocas, de regras logicamente encadeadas e de informações iniciais”, bem como “[...] tratam os dados de entrada que serão processados conforme os procedimentos definidos e geram resultados expressos em outros dados ou informações”. Mais ainda, “[...] como uma sequência de etapas bem definidas para a solução abstrata de um problema. Podemos dizer que é um conjunto de instruções finitas e encadeadas numa linguagem formal, executáveis num determinado tempo” (SILVEIRA, 2019, n.p.).

Avançando na conceituação da lógica algorítmica, e ainda pelo que apresenta Silveira (2019), considera-se relevante salientar o caráter performativo desses códigos, uma vez que, adiante, serão tratadas mais especificamente as formas como eles potencialmente constroem e interagem com perfis de usuários das redes sociais *online*. Em termos sucintos, o fato de algo ser performativo quer dizer que ele faz acontecer

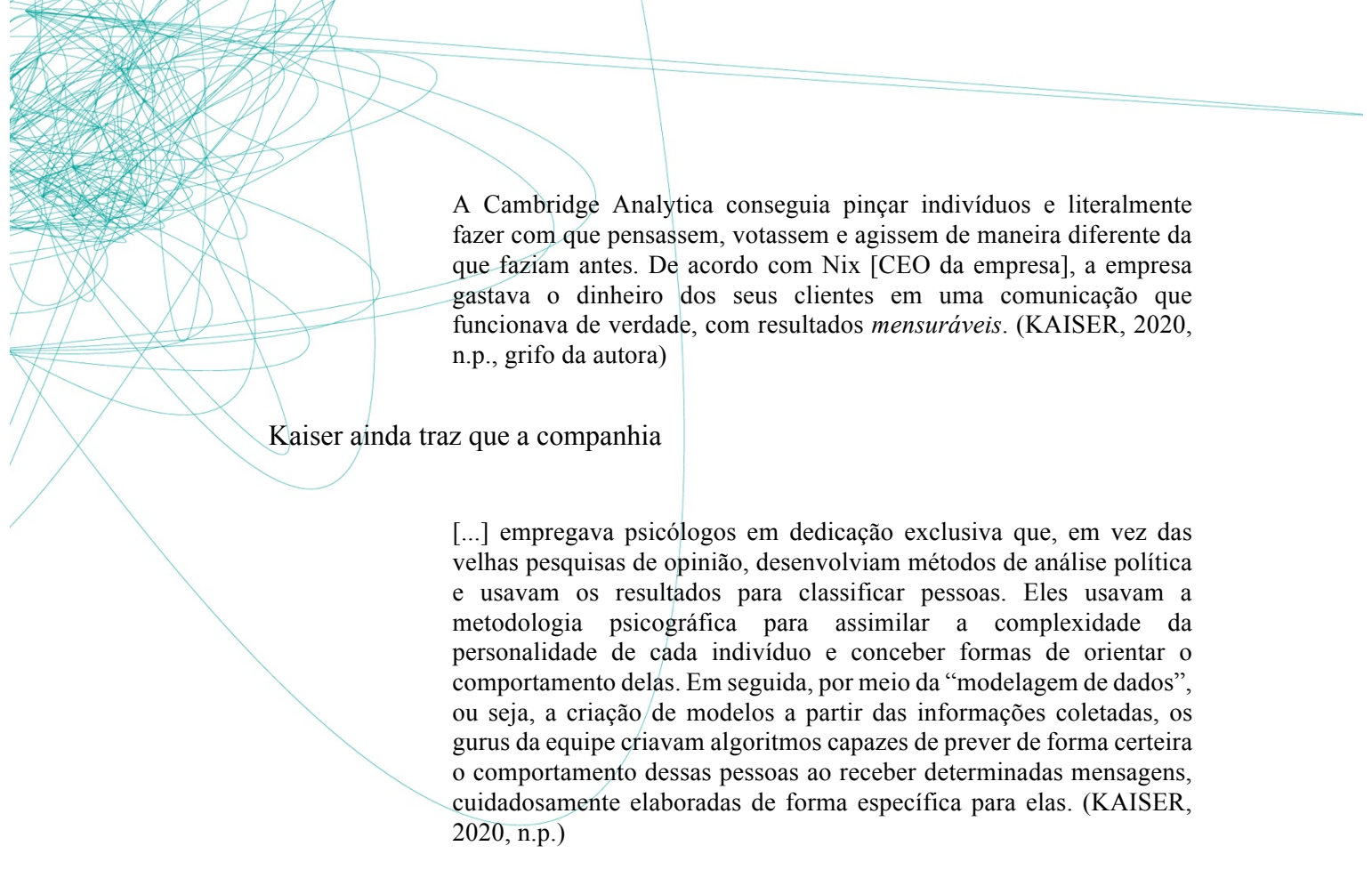
² Tradução nossa para o original em inglês: “it is only when millions and billions of individual pieces of data are linked together algorithmically that the commodity known as big data emerges.”



aquilo que narra e, por causa disso, “[...] algoritmos podem possuir viés, ou seja, um direcionamento, uma tendência e, algumas vezes, procedimentos equivocados. Como são performativos, podem promover uma série de distorções que, em inúmeros casos, podem ser apresentadas como representações fiéis e objetivas da realidade” (SILVEIRA, 2019, n.p.). Por causa disso, Jurno (2020, p. 67) acrescenta que o Facebook (nesse caso), “[...] orienta a forma que [os usuários e outros atores das redes] vão se conectar, construindo novos regimes de valor.” Tal característica performativa abrirá a possibilidade de colocar sob análise questões como predição e causação dentro do que preconiza o ponto de vista pragmaticista. A ideia de uma pretensa representação fiel ou objetiva da realidade reivindicada por sistemas algorítmicos é, também, uma outra posição de aproximação pragmaticista tanto pela impossibilidade conceitual sob a qual o signo está inserido, bem como nas formas em que seu objeto e interpretante são classificados – a saber, imediato e dinâmico.

Um dos pressupostos é que “o armazenamento de grande quantidade de dados oferece aos desenvolvedores de algoritmos a condição de inferir categorias identitárias baseadas nos hábitos de navegação na internet e no cruzamento com outras informações” (SILVEIRA, 2019, n.p), bem como o fato de que “[...] a análise de grande quantidade de dados permitir ver coisas que sentimos” (SILVEIRA, 2019, n.p). Uma possível exploração de natureza ética se dá a partir de uma espécie de violação de privacidade pelos códigos algorítmicos, porém, mais que isso, parece haver uma violação da inconsciência. Ideias como ações preemptivas baseadas em padrões extraídos de grafos psicométricos de usuários das redes, a partir de agora, não parecem pertencer a tanto a ficções distópicas assim como antes.

Avançando na descrição do objeto aqui sob análise – as estratégias da Cambridge Analytica e sua ferramenta de construção de perfis e padrões psicológicos –, cumpre-se descrevê-lo com mais detalhes. Olhando para as estratégias e ferramentas da empresa, é possível empreender questionamentos acerca do que tange o aspecto comunicacional, uma vez que ficará patente uma característica funcionalista eficiente e ineditamente mensurável, que é algo de certo modo controverso, uma vez que tal forma de olhar para o processo comunicacional foi posto em desuso e tratado como simplista, principalmente a partir da década de 1980 (RUDIGER, 2011; MARTÍN-BARBERO, 1997). Os índices de precisão preditivas da ferramenta, segundo Kaiser (2020), chegam a 95% e são baseados em estudos de cunho behaviorista. Sobre o processo da ferramenta da Cambridge Analytica:



A Cambridge Analytica conseguia pinçar indivíduos e literalmente fazer com que pensassem, votassem e agissem de maneira diferente da que faziam antes. De acordo com Nix [CEO da empresa], a empresa gastava o dinheiro dos seus clientes em uma comunicação que funcionava de verdade, com resultados *mensuráveis*. (KAISER, 2020, n.p., grifo da autora)

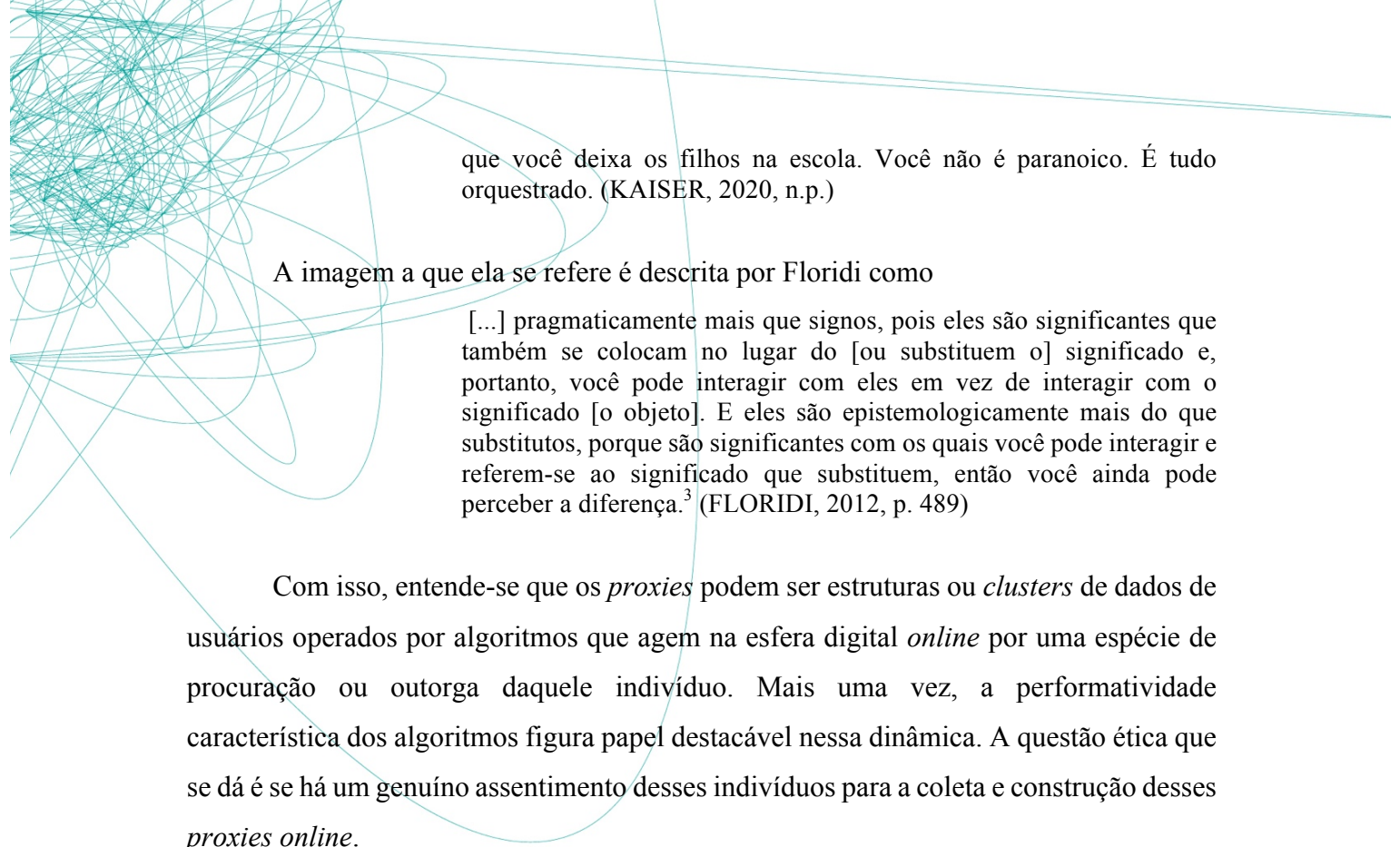
Kaiser ainda traz que a companhia

[...] empregava psicólogos em dedicação exclusiva que, em vez das velhas pesquisas de opinião, desenvolviam métodos de análise política e usavam os resultados para classificar pessoas. Eles usavam a metodologia psicográfica para assimilar a complexidade da personalidade de cada indivíduo e conceber formas de orientar o comportamento delas. Em seguida, por meio da “modelagem de dados”, ou seja, a criação de modelos a partir das informações coletadas, os gurus da equipe criavam algoritmos capazes de prever de forma certa o comportamento dessas pessoas ao receber determinadas mensagens, cuidadosamente elaboradas de forma específica para elas. (KAISER, 2020, n.p.)

Caminha-se, então na direção de maior descrição das potencialidades da ferramenta. Faz-se relevante, nesse ponto, apresentar mais um modo de aproximação da ferramenta pelo ponto de vista pragmaticista, a ser desenvolvido, que é a questão do hábito. O propósito da ferramenta da Cambridge Analytica poderia ser resumido ou simplificado em termos de mudança de hábito. Segundo Kaiser (2020), “[...] para fazer as pessoas agirem, você devia criar as condições sob as quais seria maior a probabilidade de elas fazerem o que você quer que façam. A simplicidade desse conceito era vertiginosa”. Há, portanto uma confluência dos conceitos pragmaticista de hábito, predição e causação, a serem discutidos, uma vez que hábitos parecem ser tanto a fonte a ser minerada quanto o objetivo que se pretende com a ferramenta algorítmica, em uma espécie de *loop* que corrige a própria rota.

Adiante, Kaiser traz mais uma descrição que tem, em seu bojo, a ideia de um *proxy*:

Todos esses conjuntos de dados são conhecidos como “dados comportamentais” e, com eles, os agregadores são capazes de criar uma imagem sua bastante precisa e infinitamente útil. As empresas podem, então, elaborar produtos de modo que eles se alinhem às suas atividades diárias. Os políticos usam esses dados comportamentais para lhe apresentar determinadas informações, de modo que as mensagens deles pareçam coerentes e cheguem no momento certo: pense nas propagandas sobre educação que tocam no rádio no momento exato em



que você deixa os filhos na escola. Você não é paranoico. É tudo orquestrado. (KAISER, 2020, n.p.)

A imagem a que ela se refere é descrita por Floridi como

[...] pragmaticamente mais que signos, pois eles são significantes que também se colocam no lugar do [ou substituem o] significado e, portanto, você pode interagir com eles em vez de interagir com o significado [o objeto]. E eles são epistemologicamente mais do que substitutos, porque são significantes com os quais você pode interagir e referem-se ao significado que substituem, então você ainda pode perceber a diferença.³ (FLORIDI, 2012, p. 489)

Com isso, entende-se que os *proxies* podem ser estruturas ou *clusters* de dados de usuários operados por algoritmos que agem na esfera digital *online* por uma espécie de procuração ou outorga daquele indivíduo. Mais uma vez, a performatividade característica dos algoritmos figura papel destacável nessa dinâmica. A questão ética que se dá é se há um genuíno assentimento desses indivíduos para a coleta e construção desses *proxies online*.

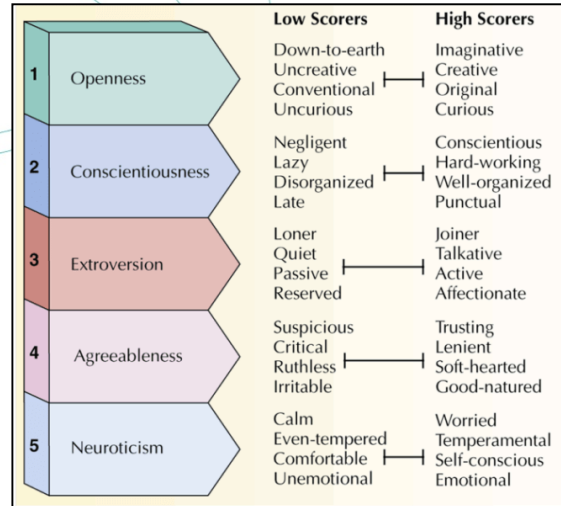
Em mais detalhes, a ferramenta da Cambridge Analytica, baseada no modelo OCEAN, é descrita por Kaiser:

[...] atribuímos a milhões de indivíduos pontuações com base no modelo “Big Five” (em geral lembrado a partir do acrônimo OCEAN), obtidas a partir dos milhares de pontos de dados que tínhamos sobre eles. A pontuação OCEAN surgiu a partir da psicologia comportamental e social no âmbito acadêmico. A Cambridge usou esse tipo de pontuação para definir de que maneira a personalidade das pessoas se constrói. Ao realizar testes de personalidade e combinar pontos de dados, a [Cambridge Analytica] descobriu que era possível determinar em que grau um indivíduo era “aberto a novas experiências” (O, de “openness”), “metódico” (C, de “conscientiousness”), “extrovertido” (E, de “extroversion”), “empático” (A, de “agreeableness”) ou “neurótico” (N, de “neuroticism”). (KAISER, 2020, n.p.)

A ferramenta OCEAN tem subvariações dentro dos 5 principais parâmetros, com intensidade de certos atributos:

³ Tradução nossa para o original “[...] proxies are pragmatically more than signs because they are signifiers that also stand in for the signified and so you can interact with them instead of interacting with the signified. And they are epistemologically more than surrogates, because they are signifiers you can interact with that refer to the signified they replace, so you can still perceive the difference.

Figura 1 – A ferramenta OCEAN e suas subdivisões, com valores mínimos e máximos



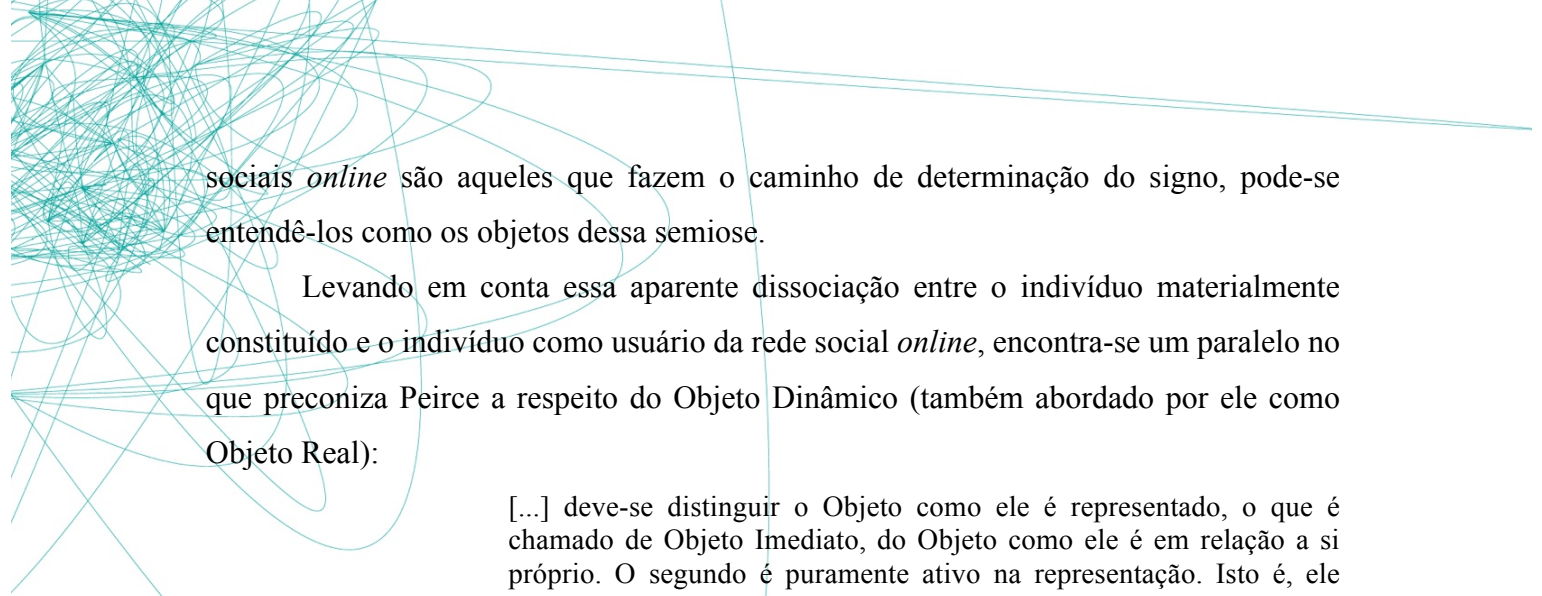
Fonte: CB Insights, 2020.⁴

Objeto, Proxy

Entende-se, pelo ponto de vista pragmaticista, que os perfis constituídos pelos elementos trazidos dentro da ferramenta OCEAN apontam indicialmente para algo, mas o que se deve levar em conta é para onde, de fato, eles apontam. Arrisca-se dizer que há uma lógica indicial vetorial em uso, uma vez que cada letra do acrônimo tem uma direção e uma intensidade. Pelos conceitos pragmaticistas, é de se levar em conta para onde ou para quem esse índice-vetor aponta – se para o indivíduo materialmente constituído, um real, um existente, ou para alguém que aquele signo **diz** apontar. Dentro da realidade do marketing, esse ente figura sob o termo *persona* ou *buyer persona*, que “[...] é um perfil fictício baseado em dados de clientes reais que representa o cliente ideal de uma empresa.”⁵ No que tange ao pragmaticismo, trata-se do (mas não se restringe ao) que se conhece como Objeto Dinâmico e Objeto Imediato. Tais perfis ou *personas* são uma espécie de generalização sob as quais pretende-se que qualquer indivíduo se encaixe. Entretanto, a ferramenta da Cambridge Analytica, segundo Kaiser (2020) permite realizar microgerenciamento das mensagens e direcionamentos de tal forma que tal generalização não aparece mais, mas sim um *targeting* hiperespecífico ou individualizado, ou melhor, *microtargeting*. O que parece é que as tanto as *buyer personas* quanto os perfis usados para *microtargeting* são constituídos de dados obtidos dos usuários (os indivíduos reais), ou seja, signos que foram coletados pela ferramenta OCEAN e sob os quais se orientam as ações de *messaging*. Considerando, com isso, que esses indivíduos que usam as redes

⁴ Disponível em <<https://www.cbinsights.com/research/what-is-psychographics/>>. Acesso em 13 jul. 2020.

⁵ Disponível em <<https://rockcontent.com/blog/personas/>>. Acesso em 13 jul. 2020.



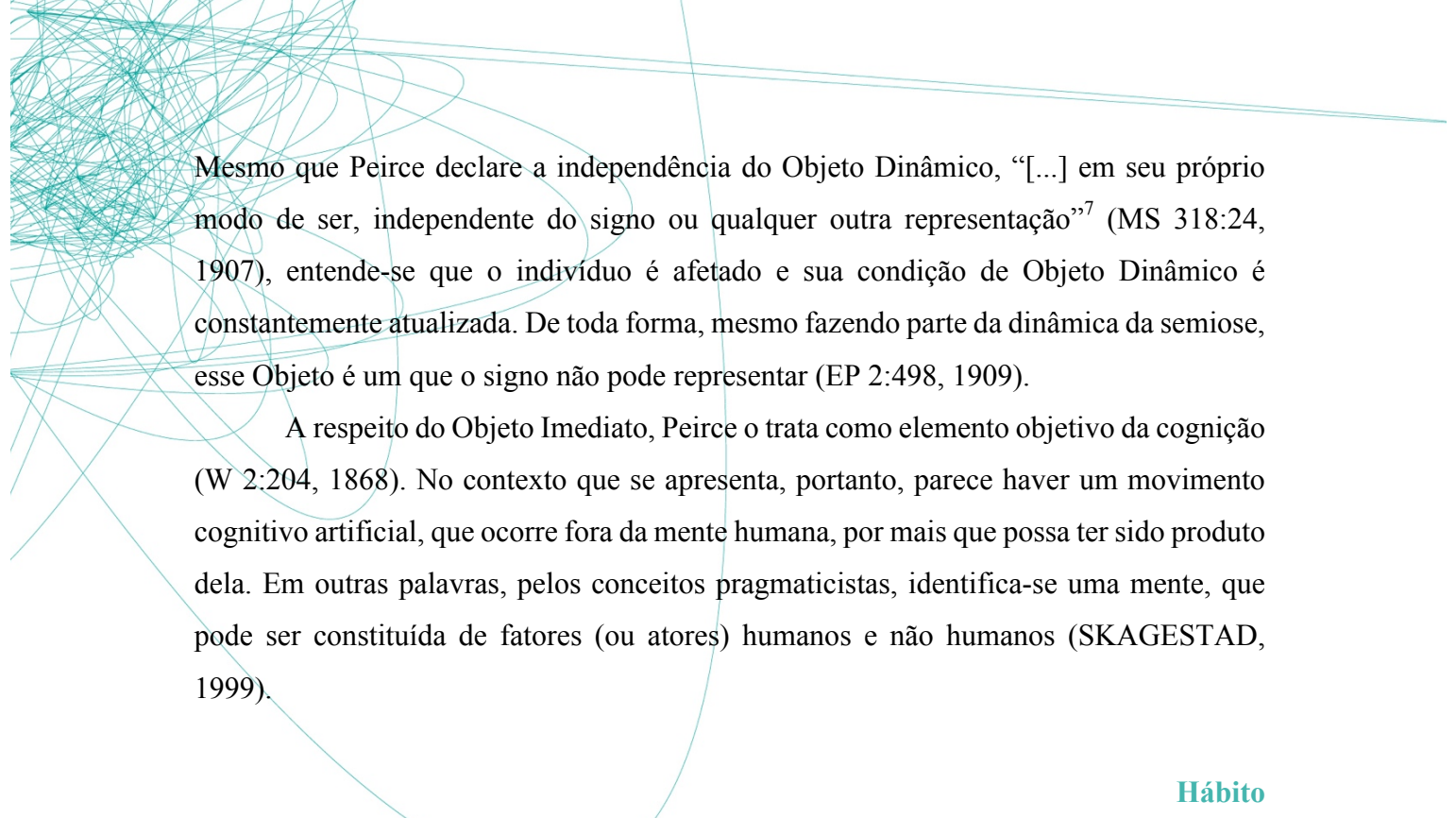
sociais *online* são aqueles que fazem o caminho de determinação do signo, pode-se entendê-los como os objetos dessa semiose.

Levando em conta essa aparente dissociação entre o indivíduo materialmente constituído e o indivíduo como usuário da rede social *online*, encontra-se um paralelo no que preconiza Peirce a respeito do Objeto Dinâmico (também abordado por ele como Objeto Real):

[...] deve-se distinguir o Objeto como ele é representado, o que é chamado de Objeto Imediato, do Objeto como ele é em relação a si próprio. O segundo é puramente ativo na representação. Isto é, ele permanece em todos os aspectos exatamente como ele era antes de ser representado. É verdade que o propósito de se representar um Objeto é geralmente, se não sempre, modificá-lo em algum aspecto. Mas o Objeto em si, o Objeto Real, é assim chamado no que refere a não ser modificado ao ser representado.⁶ (MS 793:14, 1906)

Tem-se, então, a oportunidade de identificar os tipos de Objetos propostos por Peirce. O Objeto Imediato é, como se pôde ver, aquele que faz parte ou opera na cadeia de significação; e o Objeto Dinâmico (ou Real), sendo aquele que se mantém materialmente inalterado, aquele a partir da qual iniciou-se a representação, sobre o qual a construção do Objeto Imediato apoiou-se. O curioso e conveniente ponto que se observa na fala de Peirce é a respeito da modificação do Objeto. Ora, qual não seria precisamente a intenção das estratégias da Cambridge Analytica se não modificar, de alguma forma previamente pretendida, aquele indivíduo, o Objeto Dinâmico dessa semiose, a partir os *proxies*, que aqui podem ser associados ao Objeto Imediato, ou até mesmo um conjunto deles. A partir desse ponto, é possível enxergar parte de dinâmica da semiose em questão. O Objeto Dinâmico é o ponto de partida da representação. Em seguida, pela captura de dados, produzem-se signos que apontam para aquele indivíduo que, agora, figura no *locus* de Objeto Imediato. Pelo caráter performativo dos algoritmos envolvidos, uma mudança no hábito do indivíduo é conseguida (idealmente) e um novo Objeto Dinâmico é constituído, que gera uma nova representação, por um novo Objeto Imediato. Enquanto os algoritmos e as estratégias de *messaging* funcionarem com determinado indivíduo ou enquanto esse indivíduo exibir hábitos que ligeiramente variam no tempo e que sejam detectáveis ou mensuráveis sob aqueles parâmetros, essa semiose continua em *loop*.

⁶ Tradução nossa para o original em inglês: “[...] one must distinguish the Object as it is represented, which is called the Immediate Object, from the Object as it is in itself. The latter is purely active in the representation. That is, it remains in all respects exactly as it was before it was represented. It is true that the purpose of representing an Object is usually, if not always, to modify it in some respect. But by the Object Itself, or the Real Object, we mean the Object insofar as it is not modified by being represented.”



Mesmo que Peirce declare a independência do Objeto Dinâmico, “[...] em seu próprio modo de ser, independente do signo ou qualquer outra representação”⁷ (MS 318:24, 1907), entende-se que o indivíduo é afetado e sua condição de Objeto Dinâmico é constantemente atualizada. De toda forma, mesmo fazendo parte da dinâmica da semiose, esse Objeto é um que o signo não pode representar (EP 2:498, 1909).

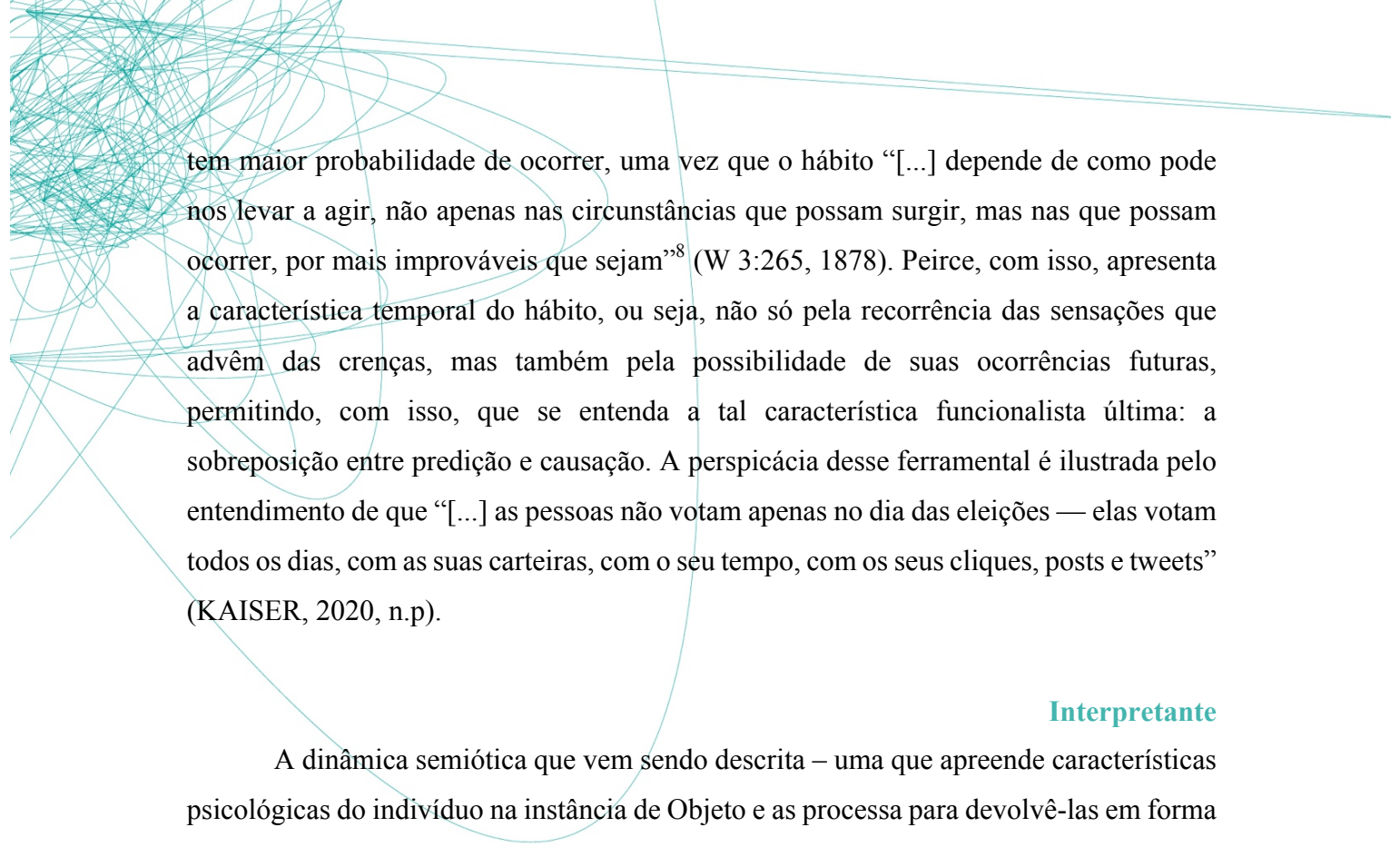
A respeito do Objeto Imediato, Peirce o trata como elemento objetivo da cognição (W 2:204, 1868). No contexto que se apresenta, portanto, parece haver um movimento cognitivo artificial, que ocorre fora da mente humana, por mais que possa ter sido produto dela. Em outras palavras, pelos conceitos pragmaticistas, identifica-se uma mente, que pode ser constituída de fatores (ou atores) humanos e não humanos (SKAGESTAD, 1999).

Hábito

O objetivo das ferramentas e das estratégias da Cambridge Analytica foi provocar uma alteração nas ações de determinados indivíduos para que eles, nos casos específicos das eleições americanas de 2016 e do referendo para a saída do Reino Unido da União Europeia no mesmo ano, tivessem uma conduta material de votar naquela diretriz pretendida. As ações de *messaging* são caracterizadas por entregar uma série de estímulos ao indivíduo, e não apenas expô-lo a postagens individuais isoladas. A repetição faz parte desse circuito uma vez que o hábito é formado por indução, para que a cognição crie a formação de uma generalidade a partir de casos isolados, gerando ações voluntárias a partir das sensações produzidas pela criação desses hábitos (W 2:232-3, 1868). A potência desse tipo de ação de *microtargeting* é dada pela sensação de verdade – ou crença – promovida pelos hábitos na cognição (W 3:265, 1878). Peirce frequentemente menciona a relação do hábito com uma fisiologia do corpo e com o sistema nervoso, reforçando o argumento apresentado a respeito da lógica de *feedback* que alimenta a dinâmica entre Objeto Dinâmico e Objeto Imediato. O hábito figura como elemento essencial nas estratégias de *microtargeting* da Cambridge Analytica pois opera em mais de um nível.

O direcionamento de anúncios, notícias ou qualquer tipo de conteúdo específicos, selecionados pela curadoria algorítmica, visam atingir aquele indivíduo repetidamente, potencialmente criando o ambiente de generalidade e sensação de verdade. Ao mesmo tempo e por causa disso eles criam, também, o contexto dentro do qual determinada ação

⁷ Tradução nossa para o original em inglês: “[...] in its own mode of being, independent of the sign or any other representation.”



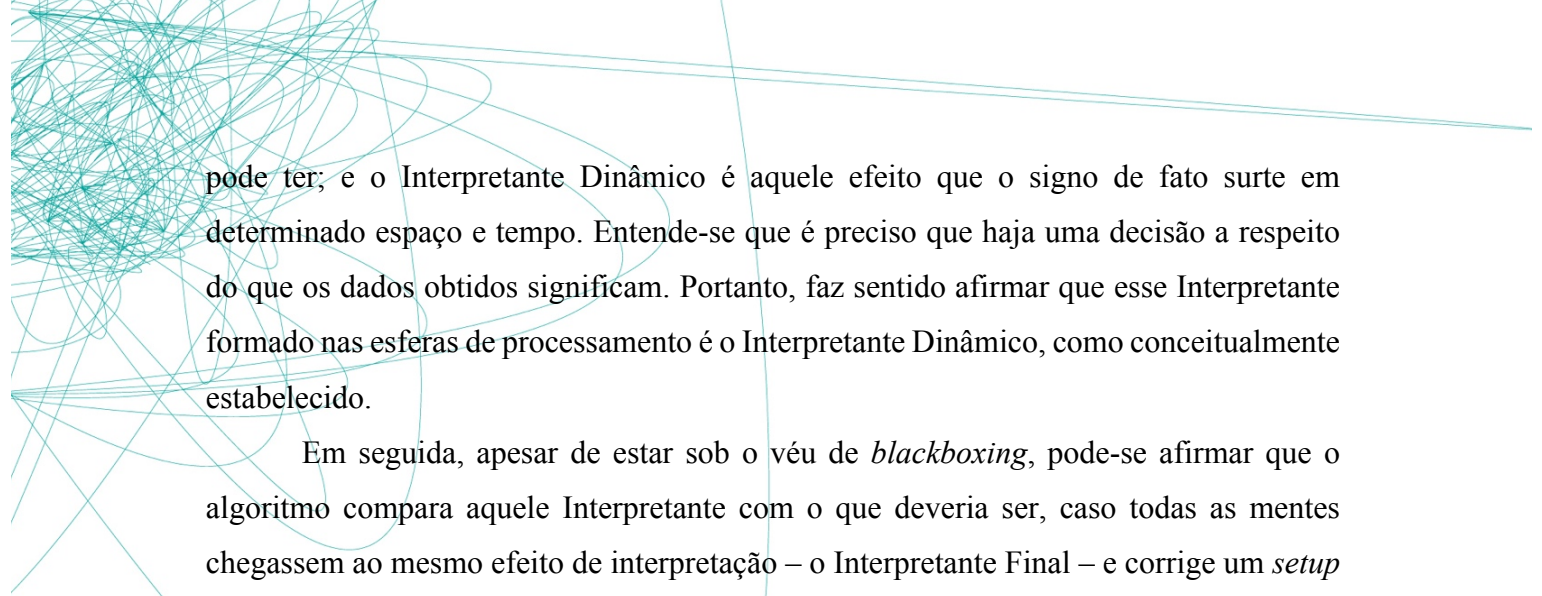
tem maior probabilidade de ocorrer, uma vez que o hábito “[...] depende de como pode nos levar a agir, não apenas nas circunstâncias que possam surgir, mas nas que possam ocorrer, por mais improváveis que sejam”⁸ (W 3:265, 1878). Peirce, com isso, apresenta a característica temporal do hábito, ou seja, não só pela recorrência das sensações que advêm das crenças, mas também pela possibilidade de suas ocorrências futuras, permitindo, com isso, que se entenda a tal característica funcionalista última: a sobreposição entre predição e causação. A perspicácia desse ferramental é ilustrada pelo entendimento de que “[...] as pessoas não votam apenas no dia das eleições — elas votam todos os dias, com as suas carteiras, com o seu tempo, com os seus cliques, posts e tweets” (KAISER, 2020, n.p).

Interpretante

A dinâmica semiótica que vem sendo descrita – uma que apreende características psicológicas do indivíduo na instância de Objeto e as processa para devolvê-las em forma de estímulos com objetivo de modulação de hábitos – opera também com a figuração dos Interpretantes. Da mesma forma que há dois lugares para o Objeto, também o Interpretante apresenta suas especificidades entre Dinâmico e Imediato. Adiante, será trazida também a ideia do Interpretante Final.

O papel do agente (ou ator) interpretador é relevante na dinâmica apontada no objeto de estudo, uma vez que, em primeiro lugar, segundo Nöth (2012), a informação só é caracterizada como tal dependendo do efeito de aumento de sentido que eventualmente surtir em sua instância de recepção; e, em segundo lugar, é também na instância de recepção que se localiza o Interpretante, possibilitando, assim, sua pauta como um dos elementos da discussão sobre os interstícios de sua relação com o Objeto por meio do signo. Salientando que, na semiose, o Interpretante é produzido ou ocorre na instância de recepção do signo, busca-se entender que recepção é essa quando se trata do processo de modulação de hábitos pelos algoritmos da Cambridge Analytica. A princípio, aponta-se que essa instância são os próprios espaços *online* de processamento, ou seja, um primeiro Interpretante aparece em uma mente não humana e ele diz respeito àquele Objeto Imediato, ou o indivíduo que o signo professa representar. Nesse sentido, o Interpretante Imediato é aquela significação em potência, ou seja, um conjunto de efeitos que o signo

⁸ Tradução nossa para o original em inglês: “[...] depends on how it might lead us to act, not merely under such circumstances as are likely to arise, but under such as might possibly occur, no matter how improbable they may be.”



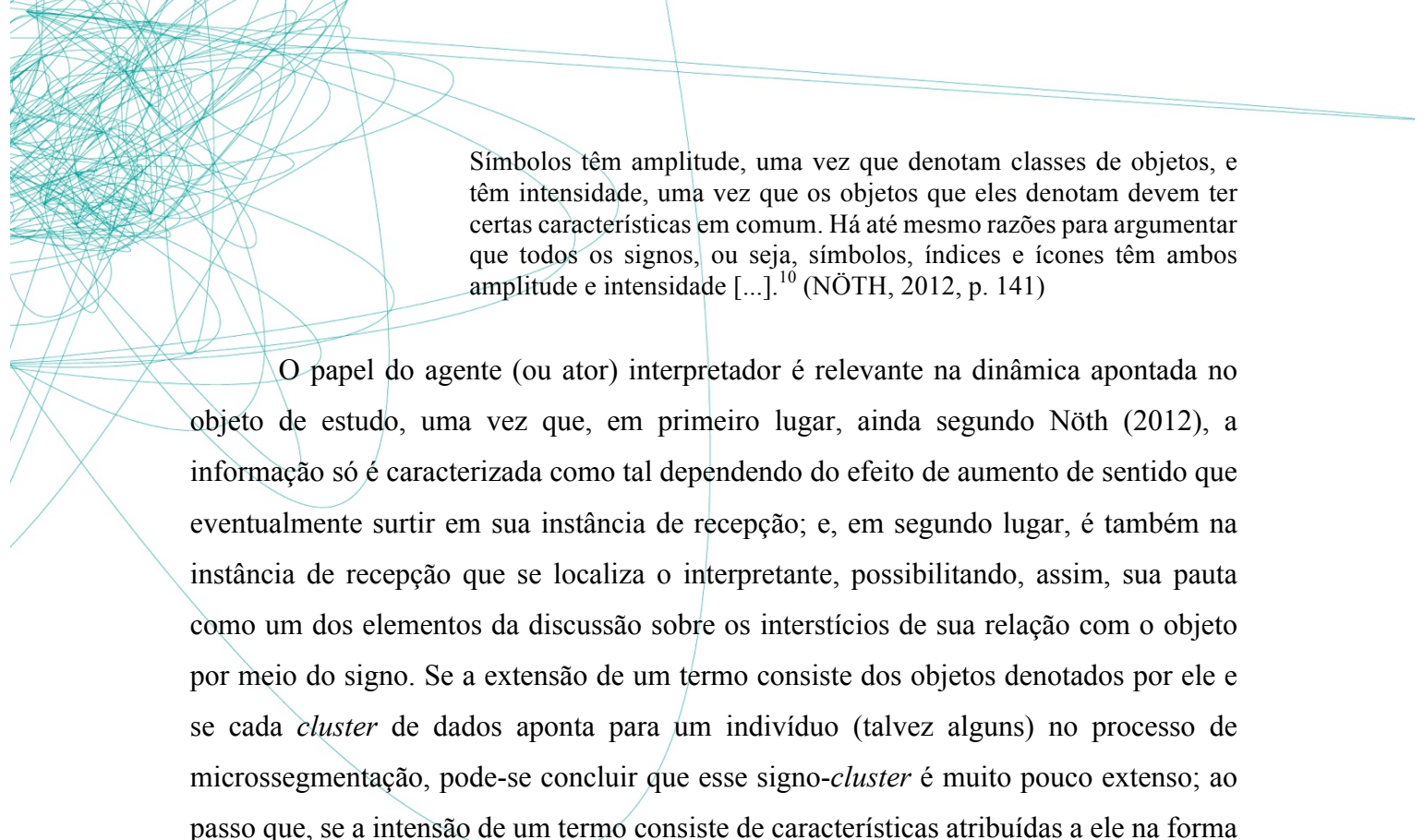
pode ter; e o Interpretante Dinâmico é aquele efeito que o signo de fato surte em determinado espaço e tempo. Entende-se que é preciso que haja uma decisão a respeito do que os dados obtidos significam. Portanto, faz sentido afirmar que esse Interpretante formado nas esferas de processamento é o Interpretante Dinâmico, como conceitualmente estabelecido.

Em seguida, apesar de estar sob o véu de *blackboxing*, pode-se afirmar que o algoritmo compara aquele Interpretante com o que deveria ser, caso todas as mentes chegassem ao mesmo efeito de interpretação – o Interpretante Final – e corrige um *setup* de estimulações no *messaging* para aquele usuário. Nesse sentido, o Interpretante Final se torna um novo signo nessa dinâmica e a nova instância de recepção, que inicialmente se localizava no ambiente *online*, agora passa a ser o indivíduo. As ferramentas de mensuração dos algoritmos, agora, irão buscar detectar com o modelo OCEAN, qual Interpretante Dinâmico foi produzido na instância da mente do indivíduo e quais potenciais cursos de ação ele poderá adotar, avaliando sucesso ou insucesso nessa iteração do processo. As conceituações do Interpretante são dadas por Santaella e Nöth (2019).

Breadth, Depth

Os algoritmos usados para processar as características detectadas no processo descrito trabalham na construção de um perfil *online* daquele usuário da rede social. Entende-se tal perfil como um conjunto de dados que fornecem uma certa informação sobre aquele indivíduo. No quadro teórico a que pertence essa discussão, aponta-se a presença do signo e da sua ação, a semiose, como o ente que coloca em circulação os sentidos que compõem um conjunto informacional, que é definido pelo “[...] valor calculado das quantidades de sua intensão e extensão”⁹ (NÖTH, 2012, p. 140). A princípio, Peirce argumenta que a composição da informação como sendo algo em relação à conjugação do intenso e extenso está estabelecida dentro da lógica do signo em seu aspecto terceiro, ou seja, seu aspecto simbólico. Como apresentou-se nessa discussão um pressuposto de que os perfis apontam indicialmente para o usuário materialmente constituído, ou esse pressuposto precisa ser colocado sob melhor e maior escrutínio, ou há também a possibilidade de o signo, em seus aspectos primeiro e segundo, também ser capaz de transportar informação:

⁹ Tradução nossa para o original em inglês: “[...] a value calculated from the quantities of its extension and intension.”



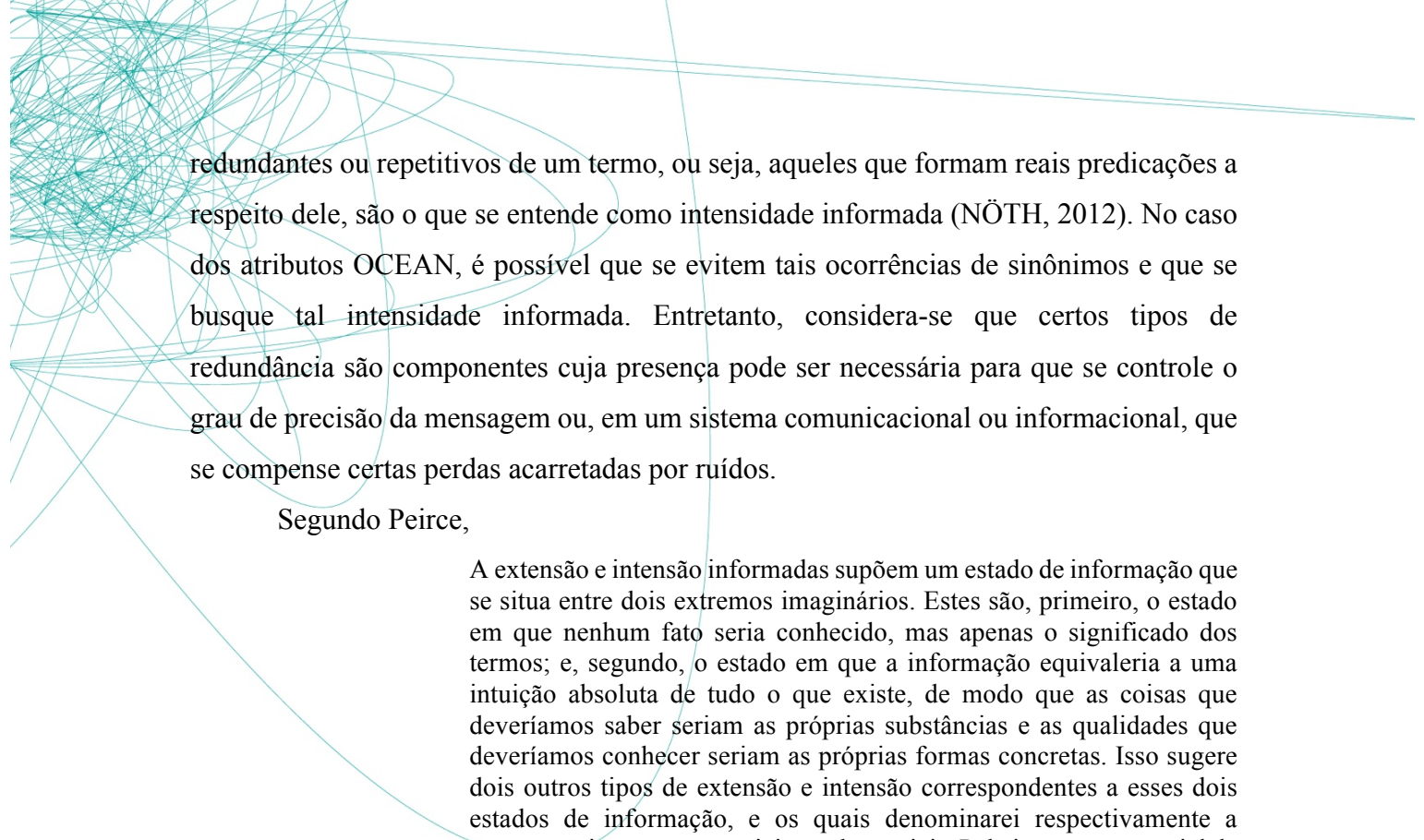
Símbolos têm amplitude, uma vez que denotam classes de objetos, e têm intensidade, uma vez que os objetos que eles denotam devem ter certas características em comum. Há até mesmo razões para argumentar que todos os signos, ou seja, símbolos, índices e ícones têm ambos amplitude e intensidade [...].¹⁰ (NÖTH, 2012, p. 141)

O papel do agente (ou ator) interpretador é relevante na dinâmica apontada no objeto de estudo, uma vez que, em primeiro lugar, ainda segundo Nöth (2012), a informação só é caracterizada como tal dependendo do efeito de aumento de sentido que eventualmente surtir em sua instância de recepção; e, em segundo lugar, é também na instância de recepção que se localiza o interpretante, possibilitando, assim, sua pauta como um dos elementos da discussão sobre os interstícios de sua relação com o objeto por meio do signo. Se a extensão de um termo consiste dos objetos denotados por ele e se cada *cluster* de dados aponta para um indivíduo (talvez alguns) no processo de microssegmentação, pode-se concluir que esse signo-*cluster* é muito pouco extenso; ao passo que, se a intensão de um termo consiste de características atribuídas a ele na forma de predicados (NÖTH, 2012), também conclui-se, nessa esteira, esse mesmo processo é muito intenso, o que condiz com a ferramenta OCEAN. Naturalmente que a extensividade reduzida é dada ao se tratar o *cluster* de dados psicométricos do indivíduo como **um** signo que aponta para ele. Entretanto, ressalva-se que a extensividade do ponto de vista do interior do *cluster* seja outra escala. Como se trata do ponto de vista de um processo de *targeting* na esfera micro, o ponto de vista da baixa extensividade faz mais sentido, tendo a quantidade de dados e parâmetros no interior desse conjunto entendida como o que confere a intensidade do termo – nesse caso, o indivíduo. Essa lógica também está de acordo com Peirce, que afirma que “quanto maior a extensão, menor é a compreensão”¹¹ (W 1:465). Com isso, desdobra-se por lógica que quanto maior a intensidade, maior é a compreensão, que é o que parece que esses dados proporcionaram aos estrategistas da Cambridge Analytica.

Peirce ainda classifica os aspectos informacionais de *breadth* e *depth* com as predicções *essencial*, *informado* e *substancial*. Como o que aparece desse processo de análise mostra que há pouca extensividade e muita intensividade, esse segundo receberá uma observação mais atenta a respeito daquelas predicções. O conjunto de atributos não

¹⁰ Tradução nossa para o original "Symbols have both extension, since they denote classes of objects, and they have intension, since the objects they denote must have certain characters in common. There are even reasons to argue that all signs, that is, symbols, indices, and icons, have both extension and intension [...]"

¹¹ Tradução nossa para o original em inglês: "The greater the extension, the less the comprehension."



redundantes ou repetitivos de um termo, ou seja, aqueles que formam reais predicções a respeito dele, são o que se entende como intensidade informada (NÖTH, 2012). No caso dos atributos OCEAN, é possível que se evitem tais ocorrências de sinônimos e que se busque tal intensidade informada. Entretanto, considera-se que certos tipos de redundância são componentes cuja presença pode ser necessária para que se controle o grau de precisão da mensagem ou, em um sistema comunicacional ou informacional, que se compense certas perdas acarretadas por ruídos.

Segundo Peirce,

A extensão e intensão informadas supõem um estado de informação que se situa entre dois extremos imaginários. Estes são, primeiro, o estado em que nenhum fato seria conhecido, mas apenas o significado dos termos; e, segundo, o estado em que a informação equivaleria a uma intuição absoluta de tudo o que existe, de modo que as coisas que deveríamos saber seriam as próprias substâncias e as qualidades que deveríamos conhecer seriam as próprias formas concretas. Isso sugere dois outros tipos de extensão e intensão correspondentes a esses dois estados de informação, e os quais denominarei respectivamente a extensão e intensão essenciais e substanciais. Pela intensão essencial de um termo, entendo as qualidades realmente concebíveis previstas em sua definição.¹² (W 2: 79-80, 1867)

Com isso, é possível propor dois lugares nessa semiose da captura de dados psicométricos dos usuários das redes sociais *online*. De fato, como apregoa Peirce, especula-se sobre esses dois lugares como sendo imaginários: em primeiro, aponta-se que o lugar onde virtualmente se localiza a intensão substancial é o inconsciente daquele indivíduo, uma vez que nesse lugar imaginário ou hipotético estariam as formas puramente intuitivas e não declaradamente predicadas; e, em segundo, o lugar onde está a leitura totalmente informatizada desse primeiro, igualmente imaginário, que é o ambiente virtual ou algum disco rígido de algum servidor, aquele que guarda os significados puros daquelas sensações puras do indivíduo.

¹² Tradução nossa para o original em inglês: “The informed breadth and depth suppose a state of information which lies somewhere between two imaginary extremes. These are, first, the state in which no fact would be known, but only the meaning of terms; and, second, the state in which the information would amount to an absolute intuition of all there is, so that the things we should know would be the very substances themselves, and the qualities we should know would be the very concrete forms themselves. This suggests two other sorts of breadth and depth corresponding to these two states of information, and which I shall term respectively the essential and the substantial breadth and depth. By the essential depth of a term, then, I mean the really conceivable qualities predicated of it in its definition.”

Conclusão

A dinâmica dessa semiose no contexto apresentado pelo documentário Privacidade Hackeada e pela autora e uma das personagens que participaram do ocorrido, Brittany Kaiser, sugere que existe uma dinâmica de um *proxy* do mundo real que se apresenta para o indivíduo em sua esfera privada, ou seja, de uma articulação de signos para que eles mostrem um mundo que se adapta a determinados interesses político-ideológicos, eleitorais, econômicos ou quaisquer outros que, porventura, possam estar em pauta. Com isso, em última instância, essas estratégias elaboradas por humanos e operadas por algoritmos buscam modular ações desses indivíduos os conduzindo por um caminho que se autocorrige e, ao se chegar em certo ponto de interesse (seja um voto, uma coleção de postagens *online* que manifestem determinado ponto de vista político ou até mesmo um hábito de consumo), buscam atuar para manter tais ações de forma recorrente, pela criação de um hábito, preferencialmente tenaz.

Referências

- ANDERSON, Chris. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. **Wired**, 2008. Disponível em: <<https://www.wired.com/2008/06/pb-theory/>>. Acesso em 15 de fevereiro de 2021.
- FLORIDI, Luciano. A *Proxy* Culture. **Philosophy & Technology**. v. 28, 2015, pp 487-490.
- JURNO, Amanda Chevtchouk. **Facebook e a plataformização do jornalismo: uma cartografia das disputas, parcerias e controvérsias entre 2014 e 2019**. 2020. 225 f. Tese (Doutorado com Comunicação Social) – FAFICH, UFMG, Belo Horizonte, 2020.
- KAISER, Brittany. **Manipulados: como a Cambridge Analytica e o Facebook invadiram a privacidade de milhões e botaram a democracia em xeque**. Rio de Janeiro: Harper Collins, 2020. n.p.
- KITCHIN, Rob. Big Data, New Epistemologies and Paradigm Shifts. **Big Data & Society**. v.1, n.1, 2014.
- KITCHIN, Rob; MCARDLE, Gavin. What Makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. **Big Data & Society**. v.3, n.1, 2016.
- MARTÍN-BARBERO, Jesus. **Dos Meios às Mediações: comunicação, cultura e hegemonia**. Rio de Janeiro: Editora da UFRJ, 1997.
- NÖTH, Winfried. Charles S. Peirce's Theory of Information: A Theory of the Growth of Symbols and of Knowledge. **Cybernetics and Human Knowing**. v. 19, n. 1-2, 2012, pp. 137-161.
- ORACLE. **What Is Big Data?** Disponível em: <https://www.oracle.com/big-data/guide/what-is-big-data.html#link2>. Acesso em 12 de setembro de 2019.

PEIRCE, Charles Sanders. **The Charles S. Peirce Papers**, 30 reels, 3rd microfilm edition. Cambridge, MA: The Houghton Library, Harvard University, Microreproduction Service (citado como MS).

_____. **The Essential Peirce**, v.2 (1893-1913), Peirce Edition Project (Ed.). Bloomington, IN: Indiana University Press (citado como EP 2).

_____. **Writings of Charles Sanders Peirce: A Chronological Edition**, ed. M. Fisch, C. Kloesel, and E. Moore, et al., 6 vols. Bloomington: Indiana University Press, 1982—. (citado como W 1, W 2 e W 3).

RUDIGER, Francisco. **As Teorias da Comunicação**. Porto Alegre: Artmed, 2011.

SANTAELLA, Lucia; NÖTH, Winfried. **Introdução à Semiótica**. São Paulo: Paulus Editora, 2019.

SILVEIRA, Sérgio Amadeu. **Democracia e os códigos invisíveis: como os algoritmos estão modulando comportamentos e escolhas políticas**. 1. ed. São Paulo: Edições SESC-SP, 2019. v. 1. e-bookp. n.p.

SKAGESTAD, Peter. Peirce's Inkstand as an External Embodiment of the Mind. **Transactions of the Charles Sanders Peirce Society**. v.35, n.3, p. 551-561. Indiana University Press, 1999.

THATCHER, Jim; O'SULLIVAN, David; MAHMOUDI, Dillon. Data Colonialism Through Accumulation by Dispossession: new metaphors for daily data. **Environment and Planning D: Society and Space**. v.34, n.6, 2016.

THE GREAT HACK. Direção de Karim Amer e Jehane Noujaim. Estados Unidos: Netflix, 2019.



semeiosis