

Avaliação da Personalidade na Era do *Big Data*: Uma Revisão Sistemática

Nathalia Melo de Carvalho*,¹

Orcid.org/0000-0001-8072-3310

Cristiane Moreira da Silva²

Orcid.org/0000-0001-8496-0233

Jean Carlos Natividade¹

Orcid.org/0000-0002-3264-9352

¹Laboratório de Pesquisa em Psicologia Social (L2PS), Departamento de Psicologia,
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brasil

²Programa de Mestrado em Cognição Social, Centro de Ciências da Saúde,
Universidade Católica de Petrópolis (UCP), Petrópolis, RJ, Brasil

Resumo

O alcance de observações naturalísticas da personalidade em ambientes on-line foi ampliado com as novas tecnologias de comunicação. Isso possibilitou capturar um grande volume de rastros digitais (*Big Data*). Esta pesquisa teve como objetivo realizar uma revisão sistemática de literatura de artigos científicos nacionais e internacionais que usaram rastros digitais para observar a manifestação da personalidade no período de 2011 a 2020. Os artigos foram buscados nas bases de dados eletrônicos PsycINFO, SciELO, Scopus (Elsevier) e PePSIC. Ao todo, foram encontrados 25 artigos, sendo que as redes sociais mais usadas para obter rastros digitais foram *Facebook* e *Twitter*. Também foi encontrado que a avaliação da personalidade por meio de rastros digitais tem sido usada com diferentes objetivos, por exemplo, para elaborar sistemas de recomendação de produtos e sistemas automáticos de detecção de *cyberbullying*. Discute-se a importância de pesquisadores brasileiros adquirirem os recursos técnicos necessários para realizar pesquisas empíricas sobre personalidade e *Big Data*, tendo em vista que apenas um dos artigos encontrados foi elaborado por pesquisadores brasileiros.

Palavras-chave: Personalidade, rastros digitais, redes sociais, métodos de observação, *Big Data*.

* Correspondência: Laboratório de Pesquisa em Psicologia Social (L2PS), Departamento de Psicologia, PUC-Rio, Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ, Brasil. CEP 22451-900. melo.nathalia@outlook.com

Personality Assessment in the Big Data Era: A Systematic Review

Abstract

The range of naturalistic observations of personality in online environments has been expanded with new communication technologies. That made it possible to capture a large volume of digital footprints (Big Data). This research aimed to carry out a systematic literature review of national and international scientific articles that used digital footprints to observe personality in the period from 2011 to 2020. The articles were accessed in the electronic databases PsycINFO, SciELO, Scopus (Elsevier), and PePSIC. In total, 25 articles were found, the most used social networks to obtain digital traces were Facebook and Twitter. It has also been found that personality assessment using digital footprints has been used for different purposes, such as developing product recommendation systems and automatic cyberbullying detection systems. The importance of Brazilian researchers to acquire the technical resources necessary to conduct empirical research on personality and big data is discussed, considering that only one of the articles found was published by Brazilian researchers.

Keywords: Personality, digital footprints, social networks, observation methods, Big Data.

Evaluación de la Personalidad en la Era de los Big Data: Una Revisión Sistemática

Resumen

La observación naturalista de la personalidad en ambientes on-line se ha ampliado con nuevas tecnologías. Esto hizo posible capturar un gran volumen de rastros digitales (*Big Data*). Esta investigación tuvo como objetivo realizar una revisión sistemática de artículos científicos nacionales e internacionales que utilizaron rastros digitales para observar la manifestación de la personalidad, en el período de 2011 a 2020. Los artículos se buscaron en las bases de datos electrónicas PsycINFO, SciELO, Scopus (Elsevier) y PePSIC. En total, se encontraron 25 artículos, las redes sociales más utilizadas para obtener huellas digitales fueron Facebook y Twitter. También se ha encontrado que la evaluación de la personalidad mediante rastros digitales se ha utilizado para diferentes propósitos, por ejemplo, para desarrollar sistemas de recomendación de productos y sistemas automáticos de detección de *cyberbullying*. Se discute la importancia de que investigadores brasileños adquieran los recursos técnicos necesarios para realizar investigaciones empíricas sobre personalidad y *big data*, considerando que solo uno de los artículos encontrados fue elaborado por investigadores brasileños.

Palabras-clave: Personalidad, pistas digitales, redes sociales, métodos de observación, Big data.

Em 2016, após o resultado das eleições presidenciais norte-americanas, as áreas de psicométrica e avaliação psicológica receberam grande notoriedade midiática. Veículos de comunicação como *The Guardian*, *New York Times* e *Vice* noticiaram o papel da ciência psicológica na vitória de Donald Trump, que contrariou expectativas e previsões de especialistas (Funk, 2016; Grassegger &

Krogerus, 2017; Hern, 2018). Uma das razões apontadas como responsáveis por esse resultado foi a estratégia de marketing inovadora da campanha eleitoral republicana. Em vez de usar dados sociodemográficos para direcionar propaganda eleitoral, Trump e sua equipe optaram por contratar a *Cambridge Analytica* – uma empresa inglesa que prometia usar grandes volumes de dados para traçar um perfil

psicológico para cada cidadão norte-americano (Grassegger & Krogerus, 2017).

O uso de rastros digitais para guiar estratégias de marketing não é algo novo no mercado publicitário. Geralmente, empresas desse segmento buscam obter informações como sexo, idade e hábitos de consumo com o objetivo de atingir o público-alvo dos clientes (Tuton & Salomon, 2015). O grande diferencial do serviço oferecido pela *Cambridge Analytica*, no entanto, consistia em acessar as características de personalidade de milhões de pessoas em poucos segundos, por meio de técnicas estatísticas e sem a necessidade de aplicar testes psicológicos (Gibney, 2018; Grassegger & Krogerus, 2017). Esses dados possibilitaram produzir conteúdos específicos para cada perfil psicológico, levando em conta características como níveis de extroversão, insegurança e instabilidade emocional (Grassegger & Krogerus, 2017).

O caso da campanha eleitoral de Trump e da *Cambridge Analytica* é apenas um exemplo de aplicação de um campo de estudo que vem crescendo na psicologia: a avaliação de construtos psicológicos por meio de rastros ou pegadas digitais. A ideia por trás dessa prática é que os comportamentos on-line das pessoas, assim como os presenciais, são manifestações de traços latentes, isto é, traços hipotéticos não observáveis ou aptidões (Kosinski et al., 2013; Primi, 2018). Assim, por exemplo, seria possível inferir os traços de personalidade das pessoas sem a necessidade de questioná-las diretamente, apenas observando a maneira como elas se comportam em ambientes on-line (Kosinski et al., 2013). Diante da atualidade e relevância do tema, esta pesquisa foi elaborada com o objetivo de realizar uma revisão sistemática de literatura de artigos científicos nacionais e internacionais que usaram rastros digitais para observar a manifestação da personalidade no período de 2011 a 2020.

Personalidade e o Modelo Big5

A personalidade, de maneira ampla, pode ser entendida como o padrão de funcionamento das pessoas no que diz respeito à maneira de

pensar, sentir e se comportar. Os estudos dessa área partem do pressuposto de que as pessoas diferem em seus aspectos intrínsecos, sendo a personalidade o conjunto de características que definem uma pessoa como ela mesma (Pervin & John, 2004). Ao longo da história da psicologia, a personalidade já foi investigada a partir de diferentes enfoques teóricos (Barenbaum & Winter, 2010). Os irmãos Floyd Allport e Gordon Allport (1921), por exemplo, se tornaram uns dos mais influentes pesquisadores dessa área ao apresentarem o conceito de traço de personalidade. Segundo G. Allport (1966), um traço representa uma tendência a responder de determinada maneira (pensar, sentir e se comportar) que é relativamente estável, isto é, sofre poucas variações decorrentes da passagem do tempo e de mudanças contextuais.

O modelo dos cinco grandes fatores, ou *Big5*, pode ser considerado uma versão moderna da teoria do traço (Natividade & Hutz, 2015). Esse modelo possui em sua concepção teórica a mesma noção de traço apresentada por G. Allport (1966), com a diferença de se tratar de um modelo empírico baseado em uma análise linguística (John et al., 1988; Natividade & Hutz, 2015). De acordo com essa hipótese, as características de personalidade podem ser identificadas na língua natural de um povo (John et al., 1988). Isso porque, ao longo do processo de socialização, as pessoas teriam inventado palavras para descrever diferenças individuais que teriam sido importantes para a sobrevivência e adaptação da espécie (e.g., Natividade & Hutz, 2016). Assim, a perspectiva lexical de estudo da personalidade defende que as diferenças individuais relevantes para descrever a espécie humana podem ser encontradas em obras de cunho literário ou léxico produzidas em determinada cultura (John et al., 1988).

Partindo dessa perspectiva, pesquisadores conduziram buscas por termos descritores da personalidade em dicionários e livros e, posteriormente, empregaram técnicas de análise fatorial para reduzir as palavras encontradas ao menor conjunto de dados possível (Goldberg,

1992). Esses estudos foram conduzidos de maneira independente nos mais diversos contextos culturais, revelando que, na maioria dos países, as características de personalidade podem ser explicadas por cinco grandes dimensões, sendo elas: extroversão, uma tendência a ser ativo e comunicativo; amabilidade, uma tendência a ser altruísta e agir de modo a atender as demandas alheias; conscienciosidade, uma tendência a ser organizado e orientado para a resolução de metas; neuroticismo, uma tendência a apresentar instabilidade emocional e emoções negativas; e abertura a experiências, uma tendência a explorar novos conhecimentos e ideias (Costa & McCrae, 2007; Hutz et al., 1998; Natividade & Hutz, 2015).

Mensuração da Personalidade

Nas últimas décadas, os cinco grandes fatores foram reconhecidos como uma estrutura capaz de descrever a personalidade de maneira consistente e replicável, o que motivou a elaboração de instrumentos de medida para acessá-los em diferentes culturas (McCrae & Costa, 2010; Pervin & John, 2004). Entre as medidas disponíveis, destacam-se as de autorrelato, as implícitas e a observação naturalística (e.g., Carvalho et al., 2022; Mehl et al., 2006; Soto & John, 2017). Cada uma dessas medidas apresenta vantagens e desvantagens, sendo mais úteis em alguns contextos do que em outros. As medidas de autorrelato, por exemplo, se destacam pela precisão e, também, pela facilidade com que podem ser inseridas em questionários de pesquisa ou em baterias de avaliação psicológica. Por outro lado, elas são mais suscetíveis à desejabilidade social (e.g., Soto & John, 2017). Já as medidas implícitas e a observação naturalística envolvem procedimentos complexos, mas apresentam as vantagens de não depender da deliberação do avaliando e capturar manifestações da personalidade mais espontâneas, sendo consideradas medidas indiretas (Asendorpf et al., 2002; Carvalho et al., 2022; Mehl et al., 2006).

Quando o interesse de um pesquisador consiste em observar como a personalidade se manifesta na vida cotidiana das pessoas, a obser-

vação naturalística é considerada a medida mais adequada (Mehl et al., 2006). Essa técnica pode ser usada tanto para observar comportamentos propriamente ditos, quanto para capturar rastros comportamentais, que seriam as pistas contextuais deixadas pelas pessoas nos ambientes em que estão naturalmente inseridas (Craig, 2000; Mehl et al., 2006). Por exemplo, pode-se inferir traços de personalidade a partir de comportamentos como manifestações de humor, ou a partir de rastros comportamentais, como a maneira como uma pessoa organiza o próprio quarto ou escritório (Gosling et al., 2002; Mehl et al., 2006). Ainda que esse método demande uma disponibilidade maior de tempo e procedimentos de análise mais robustos, é capaz de acessar a personalidade por meio da observação dos ambientes em que as pessoas estão naturalmente inseridas (Barker & Wright, 1951; Mehl et al., 2006).

Diversas tecnologias já foram usadas para acessar rastros comportamentais, assim como vários ambientes já foram alvo de observação naturalística. Nos primórdios dessa prática, os pesquisadores acompanhavam o dia a dia de pessoas e anotavam os comportamentos observados para, depois, relacioná-los com características de personalidade (Barker & Wright, 1951). Embora esse procedimento tenha sido inovador, ele foi considerado caro, trabalhoso e intrusivo, pois a presença constante de um observador podia intimidar os participantes (Mehl et al., 2006). Posteriormente, outras tecnologias surgiram, como as câmeras e os gravadores digitais, e possibilitaram observar a manifestação da personalidade de maneira mais sutil e menos custosa (Craig, 2000; Mehl et al., 2006). Até pouco tempo, entretanto, a observação naturalística era restrita a ambientes físicos. Apenas mais recentemente, por meio das novas tecnologias digitais, ambientes não-físicos se tornaram passíveis de observação (e.g., Kosinski et al., 2013).

Um dos primeiros ambientes digitais a serem usados com o propósito de observar a personalidade foram os *websites* pessoais, também conhecidos como *blogs*. Em um dos estudos pioneiros sobre esse assunto, Vazire e Gosling (2004) relacionaram rastros comportamentais

obtidos em *blogs* com escores de autorrelato de personalidade. Esse estudo mostrou que as pessoas manifestam as próprias características em páginas pessoais, principalmente no que se refere ao fator abertura a experiências. Os *blogs* foram escolhidos como tecnologias propícias à observação naturalística por serem, na época, um dos poucos ambientes digitais onde as pessoas podiam manifestar pensamentos, sentimentos e interesses (Vazire & Gosling, 2004). Atualmente, muitas outras plataformas permitem que os usuários criem conteúdo, como as redes sociais *Facebook*, *Twitter* e *Instagram*. Como consequência, estudos mais recentes ampliaram os ambientes digitais passíveis de observação naturalística, privilegiando aqueles onde as pessoas têm oportunidade para interagir e se expressar (e.g., Hagan et al., 2017; Kosinski et al., 2013).

Uma diferença entre a observação naturalística em ambientes físicos e em ambientes digitais diz respeito às tecnologias envolvidas nos procedimentos de coleta e análise de dados. Nos estudos clássicos, os dados eram coletados por meio de blocos de anotações, câmeras ou gravadores digitais – tecnologias que interferiam na rotina das pessoas, que tinham que lidar com a presença física de um observador (ainda que não humano) (Mehl et al., 2006). As novas tecnologias digitais, por outro lado, possibilitaram coletar dados de maneira automática e com o mínimo de interferência, desde que um usuário autorizasse o compartilhamento dos seus dados (Kosinski et al., 2013). Isso porque os *websites*, as redes sociais e os *blogs* já são tecnologias integradas à vida cotidiana da maioria das pessoas, sendo usadas de maneira espontânea e frequente. Além disso, as novas tecnologias superaram as anteriores no que se refere ao volume de dados armazenados (*Big Data*), que aumenta o alcance das observações naturalísticas e a capacidade de generalização dos estudos (e.g., Kosinski et al., 2013).

Personalidade e Big Data

Big Data é o termo usado para se referir a grandes volumes de dados eletrônicos indexáveis e pesquisáveis por meio de sistemas com-

putacionais (Elgendy & Elragal, 2014; Lane, 2016). Geralmente, esses dados são armazenados em servidores e analisados por meio de algoritmos, pois o grande volume de informações torna inviável interpretá-los sem o auxílio de programas específicos (Elgendy & Elragal, 2014; Lane, 2016). Esse termo também pode ser usado para se referir, mais especificamente, a uma indústria que transforma grandes volumes de dados eletrônicos em informações úteis e vendáveis (Lane, 2016). Empresas como *Google*, *Facebook* e *Twitter*, por exemplo, dispõem de dados sobre o comportamento on-line dos seus usuários, podendo vendê-los para outras empresas interessadas em obter essas informações (Lane, 2016).

O uso de *Big Data* para aferir a personalidade, no contexto da observação naturalística, teve início com as pesquisas do *The Psychometric Centre* da Universidade de Cambridge. Pela primeira vez, nesse laboratório, pesquisadores usaram grandes volumes de dados para acessar comportamentos on-line e, então, relacioná-los com características de personalidade (Kosinski et al., 2013). Esses dados foram coletados por meio de um aplicativo chamado *MyPersonality*, que foi desenvolvido com o objetivo de acessar características de personalidade de usuários do *Facebook*.

O *MyPersonality* aplicava uma bateria de testes de autorrelato de personalidade nos respondentes, que recebiam um “perfil de personalidade” baseado no modelo *Big5* e tinham a opção de compartilhar informações sobre os seus perfis digitais com a equipe de pesquisadores (Grassegger & Krogerus, 2017). O banco de dados gerado por essa coleta possibilitou avanços na área e foi usado em pesquisas que tinham diferentes objetivos, por exemplo: testar relações entre os cinco grandes fatores de personalidade e comportamentos on-line (Kosinski et al., 2013); testar um método de análise linguística de publicações digitais (Kern et al., 2014); e testar relações entre o bem-estar subjetivo e comportamentos on-line (Chen et al., 2017).

Estudos que testaram relações entre rastros digitais e traços de personalidade encontraram

padrões consistentes. Por exemplo, um estudo mostrou que pessoas extrovertidas tendem a ter mais conexões no *Facebook* do que pessoas introvertidas (Bachrach et al., 2012). Além disso, altos níveis de extroversão indicam uma probabilidade maior de uma pessoa ser fã de comediantes e atletas, enquanto pessoas altamente introvertidas tendem a consumir mais conteúdos filosóficos (Kosinski et al., 2013). Ainda, outros estudos mostraram que a linguagem usada por usuários em suas publicações também manifesta características de personalidade (e.g., Kern et al., 2014; Schwartz et al., 2013). No estudo de Kern et al. (2014), por exemplo, altos níveis de extroversão indicaram o uso de palavras e expressões como “festa”, “garotas” e “mal posso esperar”, enquanto baixos níveis nesse fator incluíram palavras representativas de atividades individuais, como “internet” e “leitura”.

Desde que esse método foi divulgado, em 2013, muitas considerações foram feitas a respeito de privacidade e ética. Os próprios autores do artigo original se mostraram preocupados com potenciais usos antiéticos da técnica criada, por exemplo, para acessar automaticamente a personalidade de indivíduos que não teriam interesse em revelá-la (Kosinski et al., 2013). Por outro lado, os autores também citaram exemplos de aplicações que seriam benéficas, como o aprimoramento de serviços digitais, considerando que as pessoas tendem a preferir conteúdos compatíveis com os seus traços de personalidade (Matz et al., 2017). Além disso, esse método pode se mostrar útil para a avaliação da personalidade em contextos em que as pessoas estão altamente motivadas a construir uma imagem favorável de si mesmas (Primi, 2018). Por se tratar de uma medida indireta, as pessoas dispõem de menos controle para omitir ou falsear respostas (Carvalho et al., 2022).

Recentemente, Primi (2018) inseriu esse tipo de medida no contexto do que denominou “testagem psicológica sem testes na era digital”, destacando-o como um dos grandes desafios da área de avaliação psicológica no século XXI. No Brasil, entretanto, a investigação desse assunto foi considerada incipiente (Limas et al., 2014).

Limas et al. (2014) mostraram que até meados desta década nenhum estudo nacional havia associado características de personalidade ao uso do *Facebook*, embora esse assunto já estivesse presente na literatura científica internacional (Kosinski et al., 2013). Uma possível explicação para esse resultado pode ser o escopo restrito da revisão de literatura, que privilegiou a rede social *Facebook*. Sabe-se que outras redes sociais como *Twitter* e *Instagram* também têm se mostrado plataformas relevantes para a investigação da personalidade (Hughes et al., 2012; Kim & Kim, 2018).

Diante da relevância do tema e da necessidade de uma revisão de literatura mais atualizada, delineou-se este estudo para responder ao seguinte problema de pesquisa: qual é o panorama científico nacional e internacional sobre a avaliação da personalidade por meio de rastros digitais? Para isso, realizou-se uma revisão sistemática de literatura em bases de dados nacionais e internacionais a fim de selecionar artigos empíricos que investigaram esse assunto no período de 2011 a 2020. Primeiro, delineou-se o escopo do material a ser buscado. Depois, definiram-se as bases de dados nas quais a busca seria realizada, assim como os descritores, os critérios de inclusão e os critérios de exclusão. Por fim, os artigos foram categorizados de acordo com critérios previamente estabelecidos.

Método

Material

O material buscado consistiu em artigos empíricos, nacionais e internacionais, que usaram rastros digitais para observar a manifestação da personalidade em redes sociais on-line no período de 01/01/2011 a 31/12/2020. Buscaram-se artigos empíricos publicados na íntegra em português, inglês ou espanhol.

Procedimentos

De Coleta e Seleção dos Estudos

A busca foi realizada em 26/01/2021 nas bases de dados eletrônicas *American Psychological Association* (PsycINFO), *Scientific*

Electronic Library Online (SciELO), Scopus (Elsevier) e Periódicos Eletrônicos de Psicologia (PePSIC). Os descritores foram selecionados com base no estudo de Azucar et al. (2018), que agrupou as palavras-chave dessa área em três grandes grupos.

O primeiro grupo consistiu de palavras usadas para se referir a redes sociais on-line: “myspace”, “facebook”, “instagram”, “twitter”, “youtube”, “photobucket”, “linkedin”, “social network” (“rede social”), “reddit”, “social media” (“mídia social”), “snapchat”, “periscope”, “social networking” (“rede social”), “status updates” (“atualização de status”), “mypersonality”. O segundo grupo consistiu de palavras usadas para se referir a diferentes procedimentos comumente usados para analisar rastros digitais: “machine learning” (“aprendizado de máquina”), “data mining” (“mineração de dados”), “text analysis” (“análise textual”), “LIWC”, “open vocabulary” (“vocabulário aberto”), “open dictionary” (“dicionário aberto”), “text mining” (“mineração de texto”), “digital footprint” (“rastros digitais”), “computational linguistics” (“linguística computacional”), “linguistic analysis” (análise linguística), “content analysis” (“análise de conteúdo”). E o terceiro grupo consistiu de palavras que se referem a traços de personalidade: “personality” (“personalidade”), “traits” (“traços”), “Big-5/Big-Five”, “Five-Factor Model” (“modelo Big5”), “extraversion” (“extroversão”), “introversion” (“introversão”), “neuroticism” (“neuroticismo”), “emotional stability” (“estabilidade emocional”), “openness” (“abertura”), “conscientiousness” (“conscienciosidade”), and “agreeableness” (“amabilidade”). As opções de busca foram preenchidas de modo que os grupos foram separados pelo recurso AND e os termos pertencentes a cada grupo, pelo recurso OR. Ainda, os campos de busca foram restringidos ao título, ao resumo e às palavras-chave de artigos científicos.

Os critérios de inclusão de estudos foram: (1) ser um artigo empírico, (2) ter associado rastros digitais ao modelo Big5, (3) ter usado redes sociais on-line para acessar automaticamente

rastros digitais, (4) ter sido publicado entre 01/01/2011 e 31/12/2020, (5) estar disponível na íntegra nas bases de dados na modalidade artigo científico, (6) ter sido publicado em inglês, espanhol ou português. Foram excluídos os artigos que usaram outras medidas que não a observação naturalística para capturar comportamentos on-line. Também foram excluídos os artigos teóricos ou de revisão.

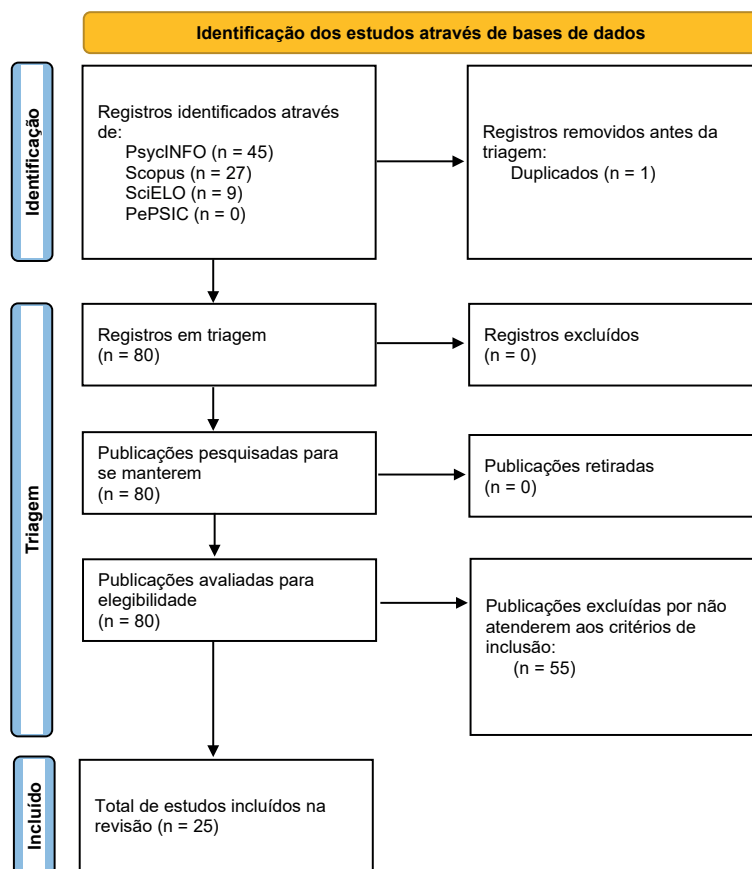
De Análises

Todos os artigos encontrados foram exportados para uma planilha do software *Microsoft Excel*. Primeiramente, investigou-se a existência de artigos duplicados. Depois, os resumos foram lidos para selecionar apenas os que respeitassem os critérios de inclusão e de exclusão. Por fim, descreveram-se os estudos encontrados de acordo com as seguintes categorias: (1) caracterização do estudo (título, autores, ano de publicação, país de origem, e periódico), (2) aspectos metodológicos (redes sociais usadas, instrumentos usados para acessar a personalidade, tamanho da amostra, procedimentos usados para analisar rastros digitais), (3) objetivo e resultados principais.

Resultados

Inicialmente, foram encontrados 81 resultados de busca, sendo que 45 foram encontrados na base de dados PsycINFO, 27 na Scopus (Elsevier) e 9 na SciELO. Não foram encontrados artigos na base PePSIC. Ainda, um artigo repetido foi localizado, restando um total de 80 artigos cujos resumos foram lidos. Após a leitura dos resumos, constatou-se que 25 artigos atendiam aos critérios de inclusão descritos acima. O fluxograma PRISMA desta revisão pode ser visualizado na Figura 1 (Page et al., 2021). Esses artigos foram, então, selecionados para análise e organizados em tabelas específicas para cada categoria de informações. Em todas as tabelas, os artigos foram organizados em ordem decrescente de ano de publicação (do mais recente ao mais antigo).

Figura 1
Fluxograma das Etapas da Pesquisa



A Tabela 1 mostra os resultados que dizem respeito à caracterização dos estudos. Destaca-se o fato de que os artigos têm origens geográficas diversas, embora haja uma prevalência de estudos de autores com afiliação institucional norte-americana ou europeia (72%). Vale ressaltar que apenas um artigo foi elaborado por autores

brasileiros (Lima & Castro, 2014). As revistas científicas onde os artigos foram publicados também são diversas, incluindo periódicos das áreas de psicologia, neurociências, ciência da computação, marketing, finanças e uma revista multidisciplinar especializada em estudos sobre interações entre humanos e animais.

Tabela 1
Caracterização dos Artigos

Título	Autores (ano)	País	Periódico
<i>Tracking fluctuations in psychological states using social media language: A case study of weekly emotion</i>	Eichstaedt e Weidman (2020)	Estados Unidos da América (EUA)	<i>European Journal of Personality</i>
<i>Differential ability of network and natural language information on social media to predict interpersonal and mental health traits</i>	Mori e Haruno (2020)	Japão	<i>Journal of Personality</i>

<i>Targeting item-level nuances leads to small but robust improvements in personality prediction from digital footprints</i>	A. N. Hall e Matz (2020)	EUA	<i>European Journal of Personality</i>
<i>Improving cyberbullying detection using Twitter users' psychological features and machine learning</i>	Balakrishnan et al. (2020)	EUA e Malásia	<i>Computers and Security</i>
<i>Does the CMO's personality matter for web traffic? Evidence from technology-based new ventures</i>	Winkler et al. (2020)	Alemanha	<i>Journal of the Academy of Marketing Science</i>
<i>Comparative analysis of feature selection algorithms for computational personality prediction from social media</i>	Marouf et al. (2020)	Bangladesh	<i>IEEE Transactions on Computational Social Systems</i>
<i>Cyberbullying detection on twitter using Big Five and Dark Triad features</i>	Balakrishnan et al. (2019)	EUA e Malásia	<i>Personality and Individual Differences</i>
<i>Mining personality traits from social messages for game recommender systems</i>	Yang e Huang (2019)	Taiwan	<i>Knowledge Based Systems</i>
<i>A personality perspective on business angel syndication</i>	Block et al. (2019)	Alemanha, Holanda e Austrália	<i>Journal of Banking and Finance</i>
<i>Facial-based personality prediction models for estimating individuals' private traits</i>	Tareaf et al. (2019)	Alemanha	<i>IEEE International Conference on Big Data and Cloud Computing (BdCloud)</i>
<i>'Who likes what and, why?' Insights into modeling users' personality based on image 'likes'</i>	Guntuku et al. (2018)	Singapura e Austrália	<i>IEEE Transactions on Affective Computing</i>
<i>Recreating the relationship between subjective wellbeing and personality using machine learning: An investigation into Facebook online behaviors</i>	Marinucci et al. (2018)	Austrália	<i>Big Data and Cognitive Computing</i>
<i>Personality profiles of users sharing animal-related content on social media</i>	Hagan et al. (2017)	EUA	<i>Anthrozoös</i>
<i>Am I who I say I am? Unobtrusive self-representation and personality recognition on Facebook</i>	M. Hall e Caton (2017)	EUA e Irlanda	<i>PLoS ONE</i>
<i>Using cognitive computing to get insights on personality traits from twitter messages</i>	Pereira e Inkpen (2017)	Canadá	<i>Advances in Artificial Intelligence</i>
<i>Analyzing Facebook activities for personality recognition</i>	Asadzadeh e Rahimi (2017)	EUA	<i>Proceedings – 16th IEEE International Conference on Machine Learning and Applications</i>
<i>PT-LDA: A latent variable model to predict personality traits of social network users</i>	Liu et al. (2016)	China	<i>Neurocomputing</i>

<i>Automatic personality assessment through social media language</i>	Park et al. (2015)	EUA e Inglaterra	<i>Journal of Personality and Social Psychology</i>
<i>Computer-based personality judgments are more accurate than those made by humans</i>	Youyou et al. (2015)	EUA e Inglaterra	<i>Proceedings of the National Academy of Sciences of the United States of America</i>
<i>Long chains or stable communities? the role of emotional stability in Twitter conversations</i>	Celli e Rossi (2015)	Itália	<i>Computational Intelligence</i>
<i>On-line consistent ranking on e-recruitment: Seeking the truth behind a well-formed CV</i>	Faliagka et al. (2014)	Grécia	<i>Artificial Intelligence Review</i>
<i>A multi-label, semi-supervised classification approach applied to personality prediction in social media</i>	Lima e Castro (2014)	Brasil	<i>Neural Networks</i>
<i>Personality, gender, and age in the language of social media: The open-vocabulary approach</i>	Schwartz et al. (2013)	EUA e Inglaterra	<i>PLoS ONE</i>
<i>Private traits and attributes are predictable from digital records of human behavior</i>	Kosinski et al. (2013)	Inglaterra	<i>Proceedings Of The National Academy Of Sciences (PNAS)</i>
<i>Be conscientious, express your sentiment!</i>	Celli e Zaga (2013)	Itália e Holanda	<i>Training</i>

A Tabela 2 mostra os aspectos metodológicos dos estudos encontrados. As redes sociais mais usadas para obter rastros digitais foram o *Facebook* e o *Twitter*. Mais especificamente, 92% dos artigos encontrados observaram a manifestação da personalidade em uma dessas duas redes. As outras redes sociais usadas foram *LinkedIn*, *Flickr* e *Steam*. Em relação aos instrumentos usados para acessar a personalidade e correlacionar com rastros digitais, houve uma prevalência de instrumentos de autorrelato, como o NEO-PI-R (Costa & McCrae, 1992) e o *International Personality Item Pool* (IPIP) (Goldberg et al., 2006). Mas também foram usados métodos indiretos ou automáticos para

acessar a personalidade, como o *IBM Watson's Personality Insights API* (International Business Machines Corporation [IBM], 2018) e o *PersonalityRecognizer* (PRec) (Mairesse et al., 2007). O tamanho das amostras dos estudos encontrados variou de 100 usuários de redes sociais on-line (Faliagka et al., 2014) a mais de 100 mil (Tareaf et al., 2019). Os procedimentos empregados nas análises, por sua vez, envolveram técnicas estatísticas derivadas dos avanços mais recentes da ciência da computação, por exemplo, *data mining* (mineração de dados), *text mining* (mineração de textos) e *machine learning* (aprendizado de máquina).

Tabela 2*Aspectos Metodológicos dos Artigos*

Autores (ano)	Redes Sociais	Medidas de Personalidade	Amostra	Análise de Rastros Digitais
Eichstaedt e Weidman (2020)	<i>Facebook</i>	<i>International Personality Item Pool (IPIP)</i> (Goldberg et al., 2006)	640 usuários do <i>Facebook</i> /303.575 atualizações de status (<i>MyPersonality</i>)	<i>Machine learning</i> ; Análise linguística – <i>Language Inquiry Word Count (LIWC)</i> (Pennebaker & Francis, 1999)
Mori e Haruno (2020)	<i>Twitter</i>	<i>Big Five</i> personality traits	239 usuários do <i>Twitter</i> (mínimo de 100 e máximo de 3.200 <i>tweets</i> por participantes)	<i>Machine Learning</i>
A. N. Hall e Matz (2020)	<i>Facebook</i>	<i>International Personality Item Pool (IPIP)</i> (Goldberg et al., 2006)	24.315 usuários do <i>Facebook</i> /31.113 likes (<i>MyPersonality</i>)	<i>Machine Learning (random forest)</i>
Balakrishnan et al. (2020)	<i>Twitter</i>	<i>IBM Watson's Personality Insights API</i> (IBM, 2018)	5.453 <i>tweets</i>	<i>Machine Learning (random forest e J48)</i>
Winkler et al. (2020)	<i>Twitter</i>	<i>Receptiviti Inc</i> (2017)	613 diretores de Marketing usuários do <i>Twitter</i> (média de 15.700 palavras por pessoa em <i>tweets</i>)	Análise linguística; <i>Language Inquiry Word Count (LIWC)</i> (Pennebaker & Francis, 1999)
Marouf et al. (2020)	<i>Facebook</i>	<i>International Personality Item Pool (IPIP)</i> (Goldberg et al., 2006)	250 usuários do <i>Facebook</i> /10 mil atualizações de status	<i>Data mining</i> ; Análise linguística - <i>Language Inquiry Word Count (LIWC)</i> (Pennebaker & Francis, 1999)
Balakrishnan et al. (2019)	<i>Twitter</i>	<i>IBM Watson's Personality Insights API</i> (IBM, 2018)	9.484 <i>tweets</i>	<i>Machine Learning (random forest)</i>
Yang e Huang (2019)	<i>Steam/Facebook</i>	<i>International Personality Item Pool (IPIP)</i> (Goldberg et al., 2006)	3 2.050 jogos/250 avaliações textuais por jogo	<i>Text Mining</i>
Block et al. (2019)	<i>Twitter</i>	<i>Receptiviti Inc</i> (2017)	1.348 <i>business angles</i> usuários do <i>Twitter</i> (média de 2.566 <i>tweets</i>)	Análise linguística; <i>Language Inquiry Word Count (LIWC)</i>

Tareaf et al. (2019)	<i>Facebook/ Twitter</i>	NEO-PI-R (Costa & McCrae, 2008)	108.547 usuários do <i>Facebook</i> / 610 usuários do <i>Twitter</i> (média de 2800 <i>tweets</i> por usuário)	Análise linguística: <i>Language Inquiry Word Count</i> (LIWC) (Pennebaker & Francis, 1999); Análise facial: <i>Face++ tool</i> ; <i>Machine Learning</i>
Guntuku et al. (2018)	<i>Flickr</i>	BFI-10 (Rammstedt & John, 2007)	300 usuários do <i>Flickr</i> com 200 <i>likes</i> cada (60.000 imagens)	<i>Machine Learning</i>
Marinucci et al. (2018)	<i>Facebook</i>	NEO-PI-R (Costa & McCrae, 2008)	21.122 usuários do <i>Facebook</i> (<i>MyPersonality</i>)	<i>Machine Learning</i>
Hagan et al. (2017)	<i>Facebook/ Twitter</i>	NEO-PI-R (Costa & McCrae, 2008)	Estudo 1: 72.559 usuários do <i>Facebook</i> / 14.094.365 atualizações de status (<i>MyPersonality</i>). Estudo 2: 3.810 usuários do <i>Twitter</i> / 663.469 imagens	Análise linguística; <i>Machine Learning</i>
M. Hall e Caton (2017)	<i>Facebook</i>	<i>Big Five Inventory</i> (John et al., 1991)	509 usuários do <i>Facebook</i>	Análise linguística; <i>Language Inquiry Word Count</i> (LIWC) (Pennebaker & Francis, 1999)
Pereira e Inkpen (2017)	<i>Twitter</i>	<i>IBM Watson's Personality Insights API</i> (IBM, 2018)	152 usuários do <i>Twitter</i> (100 <i>tweets</i> por usuário)	<i>Text mining</i> ; <i>Machine Learning</i>
Asadzadeh e Rahimi (2017)	<i>Facebook</i>	<i>Big Five personality test</i>	92.225 usuários do <i>Facebook</i> (<i>MyPersonality</i>)	<i>Machine Learning</i>
Liu et al. (2016)	<i>Facebook</i>	NEO-PI-R (Costa & McCrae, 1992)	5.135 usuários do <i>Facebook</i> / Mais de 10 milhões de palavras (<i>MyPersonality</i>)	<i>Machine Learning</i>
Park et al. (2015)	<i>Facebook</i>	NEO-PI-R (Costa & McCrae, 1992)	71.556 usuários do <i>Facebook</i> /15 milhões de atualizações de <i>status</i>	Análise linguística; <i>Open-vocabulary approach</i>
Youyou et al. (2015)	<i>Facebook</i>	NEO-PI-R (Costa & McCrae, 1992)	86.220 usuários do <i>Facebook</i> com pelo menos 20 <i>likes</i> cada (mais de 1 milhão e 500 <i>likes</i>)	<i>Data mining</i>
Celli e Rossi (2015)	<i>Twitter</i>	<i>Personage</i> (Mairesse & Walker, 2007); <i>Analyzeweords</i> (Tausczic & Pennebaker, 2010)	Mais de 13.000 usuários do <i>Twitter</i> / 200.000 <i>tweets</i>	Análise linguística

Faliagka et al. (2014)	<i>LinkedIn</i>	Níveis de extroversão foram extraídos da quantidade de palavras representativas de atividades sociais e de emoções positivas em publicações on-line	100 usuários do <i>LinkedIn</i> que aplicaram para uma vaga de emprego	Análise linguística; <i>Language Inquiry Word Count</i> (LIWC) (Pennebaker & Francis, 1999)
Lima e Castro (2014)	<i>Twitter</i>	<i>Personality Recognizer</i> (PRec) (Mairesse et al., 2007)	18.435 <i>tweets</i> (obtidos em três bases de dados diferentes)	<i>Machine Learning</i>
Schwartz et al. (2013)	<i>Facebook</i>	NEO-PI-R (Costa & McCrae, 1992)	74.941 usuários / Mais de 15 milhões de atualização de <i>status</i> (<i>MyPersonality</i>)	Análise linguística; <i>Language Inquiry Word Count</i> (LIWC) (Pennebaker & Francis, 1999); <i>Open-vocabulary approach</i>
Kosinski et al. (2013)	<i>Facebook</i>	<i>International Personality Item Pool</i> (IPIP) (Goldberg et al., 2006)	58.466 usuários do <i>Facebook</i> /170 likes por usuário, em média (<i>MyPersonality</i>)	<i>Data mining</i>
Celli e Zaga (2013)	<i>Twitter</i>	BFI-10 (Rammstedt & John, 2007)	5.816 usuários do <i>Twitter</i> (média de 12 <i>tweets</i> por usuário)	Análise linguística

Por fim, a Tabela 3 mostra o objetivo e os principais resultados dos artigos encontrados. Destacam-se quatro eixos temáticos: (1) estudos que buscaram identificar padrões de relações entre rastros digitais e traços de personalidade (e.g., Hagan et al., 2017; Kosinski et al., 2013); (2) estudos que usaram a observação naturalística em ambientes digitais como instrumento de medida para acessar traços de personalidade e, então, relacioná-los com outras variáveis (e.g., Block et al., 2019; Celli & Zaga, 2013); (3) estudos que apresentaram avanços técnicos, por exemplo, novos procedimentos de coleta e análise

de rastros digitais (e.g., Liu et al., 2016; Marouf et al., 2020); (4) e estudos que testaram sistemas automatizados de recomendação ou detecção baseados em rastros digitais (e.g., um sistema de detecção de *cyberbullying*, Balakrishnan et al., 2020; um sistema de recomendação de jogos, Yang & Huang, 2019). De forma geral, os resultados dos estudos apontaram para a direção de ser possível prever a personalidade por meio de rastros digitais e de que essa forma de medida contribui para a automação de processos que antes eram manuais.

Tabela 3*Objetivo e Resultados Principais dos Artigos*

Autores (ano)	Objetivo	Resultados Principais
Eichstaedt e Weidman (2020)	Testar o potencial de um método baseado em big data para ajudar a rastrear flutuações intrapessoais em estados dinâmicos de personalidade	Rastros digitais podem ser usados para acessar a experiência emocional das pessoas, ao mesmo tempo que possibilita estabelecer relações entre estados emocionais e as principais dimensões da personalidade (<i>Big5</i>)
Mori e Haruno (2020)	Investigar em que medida diferentes informações coletadas no <i>Twitter</i> (e.g., número de tweets e de conexões) predizem traços para além do modelo <i>Big5</i> (e.g., ansiedade; maquiavelismo)	O uso do <i>Twitter</i> prediz traços de personalidade, por exemplo, a linguagem prediz traços relacionados a saúde mental (e.g., ansiedade e esquizofrenia) e a rede de conexões prediz traços relacionados a interação social (e.g., autismo)
A. N. Hall e Matz (2020)	Testar o poder preditivo de rastros digitais sobre as nuances (itens específicos) da personalidade (<i>Big5</i>)	Uma análise preditiva baseada nas nuances da personalidade (nas respostas aos itens, ao invés das médias dos fatores) aumenta o poder preditivo de rastros digitais sobre traços
Balakrishnan et al. (2020)	Testar um sistema automático de detecção de <i>cyberbullying</i> baseado na personalidade (<i>Big5</i>), sentimentos e emoções	A detecção de <i>cyberbullying</i> se torna mais acurada quando a personalidade e os sentimentos são acessados por meio de rastros digitais
Winkler et al. (2020)	Investigar se a personalidade dos Diretores de Marketing (CMOs) de novos empreendimentos baseados em tecnologia afeta como a crescente maturidade de novos empreendimentos se traduz em tráfego na web	A extroversão de um CMO modera positivamente a relação entre a maturidade de um novo empreendimento e o tráfego da <i>Web</i> , enquanto a conscienciosidade de um CMO é um moderador negativo dessa relação
Marouf et al. (2020)	Comparar o desempenho de cinco algoritmos de seleção de recursos (<i>machine learning</i>) no contexto da avaliação da personalidade por meio de rastros digitais	O algoritmo de seleção de recursos baseado no coeficiente de correlação de Pearson (PCC) proporciona uma acurácia maior na predição da personalidade do que outros algoritmos usados na literatura
Balakrishnan et al. (2019)	Testar um modelo de detecção de <i>cyberbullying</i> baseado no <i>Big5</i> e na <i>Dark Triad</i>	A avaliação dos fatores extroversão, amabilidade, neuroticismo e psicopatia por meio de rastros digitais aumenta a precisão de um modelo para identificar <i>cyberbullying</i>
Yang e Huang (2019)	Testar um sistema de recomendação de jogos baseado no modelo <i>Big5</i>	Jogadores se mostram satisfeitos com o sistema de recomendação baseado em uma consonância entre a “personalidade” dos jogos e a personalidade dos jogadores

Block et al. (2019)	Testar o papel da personalidade (<i>Big5</i>) nas decisões de investir em novos negócios individualmente ou em conjunto com um grupo de investidores (<i>syndication</i>)	Altos níveis de extroversão tornam o investimento em grupo mais provável, enquanto altos níveis de conscienciosidade diminuem a probabilidade dessa forma de investimento
Tareaf et al. (2019)	Investigar como as imagens de perfil do <i>Twitter</i> diferem com base na personalidade dos usuários que as publicam	Existem variações nas fotos de perfil de pessoas com diferentes traços de personalidade. Por exemplo, mulheres com alta abertura a experiências apresentam uma tendência a publicar fotos de perfil que indicam emoções positivas (e.g., sorriso aparente)
Guntuku et al. (2018)	Testar o poder preditivo de <i>likes</i> em imagens sobre a personalidade (<i>Big5</i>)	Fotos de perfil predizem negativamente o fator abertura a experiências e conteúdos relacionados a desordem predizem positivamente o fator neuroticismo
Marinucci et al. (2018)	Aprofundar a compreensão da relação entre o bem-estar subjetivo e o modelo <i>Big5</i> , tentando replicar a relação usando modelos de previsão de aprendizado de máquina baseados em likes do <i>Facebook</i>	As relações entre bem-estar subjetivo e extroversão, neuroticismo e conscienciosidade foram replicadas com sucesso no modelo de aprendizado de máquina.
Hagan et al. (2017)	Testar relações entre traços de personalidade (<i>Big5</i>) e preferência por animais (cães e gatos) manifestada em redes sociais on-line	A presença de cães em imagens está associada a maiores níveis de extroversão, enquanto a presença de gatos está associada a maiores níveis de abertura
M. Hall e Caton (2017)	Testar o efeito do manejo de impressões na avaliação da personalidade (<i>Big5</i>) por meio de rastros digitais	Rastros digitais textuais predizem os traços de personalidade, mas são apresentadas evidências de manejo de impressões no uso do <i>Facebook</i>
Pereira e Inkpen (2017)	Testar um modelo de predição da personalidade (<i>Big5</i>) baseado em análise linguística e <i>Deep Learning</i>	Uma técnica de análise linguística baseada em <i>Deep Learning</i> melhora a predição de traços de personalidade
Asadzadeh e Rahimi (2017)	Testar o poder preditivo de rastros digitais sobre os cinco grandes fatores de personalidade por meio do algoritmo LASSO	O melhor nível de precisão é alcançado para a predição dos fatores abertura e extroversão; o nível de precisão mais baixo é obtido para amabilidade
Liu et al. (2016)	Testar um novo modelo probabilístico (PT-LDA model) para prever a personalidade por meio de rastros digitais	O modelo PT-LDA se mostra mais acurado, eficiente e robusto do que outros modelos disponíveis na literatura
Park et al. (2015)	Testar um método para acessar a personalidade por meio de uma análise de vocabulário aberto (<i>open-vocabulary approach</i>) obtido em redes sociais on-line	Uma análise da linguagem obtida em redes sociais se mostra uma medida precisa e com evidências de validade para acessar a personalidade (relações esperadas com outras variáveis e estabilidade temporal satisfatória)

Youyou et al. (2015)	Comparar a acurácia de teorias implícitas da personalidade com a acurácia de modelos computacionais de predição da personalidade	Um modelo computacional de predição da personalidade baseado em rastros digitais do <i>Facebook</i> se mostra mais preciso do que as teorias implícitas da personalidade (avaliação por pares)
Celli e Rossi (2015)	Investigar como a estabilidade emocional afeta as relações sociais no <i>Twitter</i> por meio de um novo software de análise linguística	Usuários seguros tendem a construir redes mais fortes, enquanto usuários neuróticos têm mais dificuldade em pertencer a uma comunidade estável; dessa forma, eles buscam por mais contatos novos em redes sociais
Faliagka et al. (2014)	Testar um sistema para avaliar automaticamente o perfil de candidatos a vagas de emprego por meio de <i>machine learning</i>	O sistema tem um desempenho consistente em comparação com recrutadores humanos, portanto, pode ser confiável para a automação da classificação de candidatos e mineração da personalidade
Lima e Castro (2014)	Testar um sistema de predição da personalidade (PERSOMA) por meio de rastros digitais baseado em grupos de palavras, ao invés de textos únicos	O sistema apresenta uma acurácia aproximada de 83%, sendo que os fatores extroversão, amabilidade e neuroticismo são mais bem preditos do que os outros fatores
Schwartz et al. (2013)	Testar relações entre traços de personalidade (<i>Big5</i>) e o uso de palavras/expressões em atualizações de status no <i>Facebook</i>	Altos níveis de neuroticismo indicam uma tendência a usar expressões/palavras como “de saco cheio” ou “deprimido” e altos níveis de estabilidade emocional indicam uma tendência a usar palavras que se referem a uma vida ativa (e.g., esportes)
Kosinski et al. (2013)	Testar o poder preditivo de <i>likes</i> de usuários do <i>Facebook</i> sobre traços de personalidade (<i>Big5</i>)	<i>Likes</i> predizem traços de personalidade nas seguintes direções: atividades sociais predizem o fator extroversão; afetos negativos, o fator neuroticismo; conteúdos artísticos, o fator abertura; conteúdos religiosos, o fator amabilidade; conteúdos de trabalho, o fator conscienciosidade
Celli e Zaga (2013)	Testar o poder preditivo de traços de personalidade e de algumas pistas linguísticas, incluindo <i>hashtags</i> , para prever a subjetividade e a polaridade de sentimentos no <i>Twitter</i>	A posição de <i>hashtags</i> e os níveis de conscienciosidade se mostraram preditores de manifestações de sentimentos no <i>Twitter</i>

Discussão

Diante dos resultados encontrados, considera-se que o panorama nacional ainda é incipiente, uma vez que apenas um artigo

encontrado foi elaborado por pesquisadores brasileiros (Lima & Castro, 2014). Esse resultado vai ao encontro do estudo de Lima et al. (2014), que apontou para uma ausência de artigos nacionais que associaram características de

personalidade ao uso do *Facebook*. Ao contrário do estudo de Limas et al. (2014), porém, nesta revisão buscaram-se artigos que usaram uma ampla gama de redes sociais on-line, para além do *Facebook*, como *Twitter*, *Instagram* e *LinkedIn*. Essa busca mais completa possibilitou encontrar um maior número de estudos e revelou que o *Twitter* também tem se mostrado uma rede social relevante nas investigações dessa área (e.g., Balakrishnan et al., 2019; Block et al., 2019). O único estudo realizado por pesquisadores brasileiros, por exemplo, usou uma amostra de usuários do *Twitter* para testar um modelo de predição de traços de personalidade por meio de rastros digitais (PERSOMA). Esse modelo difere de outros apresentados na literatura pois trabalha com grupos de textos e não usa diretamente o conteúdo das mensagens, extraindo metaatributos dos textos (Lima & Castro, 2014).

Uma análise mais detalhada dos resultados encontrados nesta pesquisa, principalmente dos aspectos metodológicos dos artigos, permite a elaboração de algumas hipóteses que podem estar na base da escassez de estudos empíricos no contexto brasileiro. A análise de *Big Data* envolve procedimentos de análise complexos (e.g., *data mining*; *text mining*; *machine learning*), que não são abarcados em cursos de graduação e pós-graduação em psicologia. Nesse sentido, por mais que haja interesse teórico de pesquisadores brasileiros nessa área (e.g., Limas et al., 2014; Primi, 2018), a ausência de instrumentalização técnica pode resultar em um número pequeno de profissionais capacitados para realizar pesquisas empíricas sobre personalidade e *Big Data*.

Embora a observação naturalística da personalidade em ambientes digitais remeta ao início dos anos 2000, o uso de *Big Data* nessa área ainda é considerado uma tecnologia nova, tendo sido impulsionada por avanços recentes da ciência da computação (Kosinski et al., 2013). Nesse sentido, pode-se compreender a multidisciplinaridade das revistas onde os artigos dessa área são publicados. Os artigos encontrados foram publicados em revistas especializadas em diversos assuntos, como inteligência artificial

(e.g., Faliagka et al., 2014), marketing (e.g., Winkler et al., 2020) e psicologia (e.g., Park et al., 2015). Esse resultado evidencia uma possível articulação entre diversas áreas do conhecimento para que seja possível observar a manifestação da personalidade por meio de grandes volumes de dados. Um treinamento mais robusto em análise de rastros digitais para pesquisadores brasileiros e uma maior articulação da psicologia com outros departamentos, como o de ciência da computação, pode resultar em mais estudos sobre esse assunto.

No que diz respeito aos instrumentos usados pelos estudos encontrados para relacionar rastros digitais com características de personalidade, pode-se observar uma prevalência de instrumentos de autorrelato. Possivelmente, esses instrumentos foram usados por serem precisos e devido à facilidade com que podem ser inseridos em questionários de pesquisa. Assim, por exemplo, o aplicativo *MyPersonality* podia aplicar facilmente medidas como o NEO-PI-R (Costa & McCrae, 1992) e o IPIP (Goldberg et al., 2006) nos participantes, ao mesmo tempo em que capturava informações automaticamente dos perfis digitais. Outros estudos, ainda, usaram modelos computacionais de predição de características psicológicas, partindo do pressuposto de que essa já é uma forma de medida precisa e com evidências de validade (e.g., Balakrishnan et al., 2019). Não foram encontrados, porém, estudos que usaram medidas implícitas de conteúdos cognitivos. Estudos futuros podem associar, por exemplo, testes de associação implícitas (IAT) (e.g., Carvalho et al., 2022) à observação naturalística da personalidade, uma vez ambas as formas de medida podem ser consideradas indiretas.

Amostras obtidas por meio do aplicativo *MyPersonality* foram recorrentes nos estudos encontrados. Pode-se observar, no entanto, que a obtenção de rastros digitais não se resume a esse aplicativo. Outros estudos capturaram um grande número de observações por meio de outros métodos (e.g., Balakrishnan et al., 2020). Ainda, alguns estudos usaram amostras mais específicas e menores, por exemplo, Guntuku

et al. (2018), que acessaram os perfis digitais de 300 usuários do *Flickr*. Vale ressaltar que mesmo estudos com um número reduzido de participantes podem ser inseridos no contexto do *Big Data*, uma vez que cada usuário de redes sociais on-line produz muitas variáveis. Assim, por exemplo, Guntuku et al. (2018) obtiveram 60.000 observações mesmo com apenas 300 participantes, pois cada um deles havia “curtido” pelo menos 200 imagens.

Tendo em vista os quatro eixos temáticos dos estudos encontrados, considera-se que muitos avanços têm sido feitos nessa área, após o estudo pioneiro de Kosinski et al. (2013). Todos os estudos encontrados apresentaram resultados favoráveis à predição da personalidade por meio de rastros digitais, pelo menos em algum nível. Uma metanálise recente ajudou a elucidar o tamanho de efeito dessa predição, Azucar et al. (2018) verificaram que os resultados dos estudos dessa área estão de acordo com o que se espera de comportamentos predizendo traços, seja em ambientes físicos ou digitais. Nessa metanálise, as correlações entre rastros digitais e traços de personalidade variaram de 0,29, para o fator amabilidade, a 0,40, para o fator extroversão.

No que diz respeito a resultados específicos dos estudos encontrados, estudos mais recentes se preocuparam em associar traços de personalidade obtidos por meio de rastros digitais com comportamentos mais específicos. Balakrishnan et al. (2019), por exemplo, mostraram que maiores níveis de psicopatia e menores níveis de amabilidade estão associados a uma tendência a praticar *cyberbullying*. Desse modo, a avaliação de características que compõem o *Big5* e a *Dark Triad*, que se refere a traços sombrios da personalidade (psicopatia, narcisismo e maquiavelismo), por meio de rastros digitais pode resultar em um sistema automático para capturar *cyberbullying* (Balakrishnan et al., 2019). Ainda, Faliagka et al. (2014) mostraram que é possível acessar características de personalidade de usuários do *LinkedIn* que se candidatam a vagas de emprego sem o uso de testes psicológicos, facilitando a automação e, conseqüentemente, a agilidade de processos seletivos. Isso não signi-

fica que os rastros digitais substituirão os testes psicológicos, mas que podem complementá-los, sem desconsiderar os vieses e fontes de erro inerentes a qualquer medida em psicologia.

Esses resultados apontam para a direção de que a personalidade pode, sim, ser observada em ambientes digitais e que os padrões encontrados podem ser usados para fins aplicados. A esse respeito, vale ressaltar que a avaliação da personalidade por meio de rastros digitais ganhou notoriedade midiática por ter sido usada para fins eleitorais, mais especificamente, para a elaboração de mensagens persuasivas consonantes com determinados perfis psicológicos (Grassegger & Krogerus, 2017). No entanto, sistemas de recomendação de candidatos/produtos não são a única possibilidade de aplicação dessa técnica. Um sistema de detecção de traços que compõem a *Dark Triad*, por exemplo, pode facilitar a identificação de crimes digitais. Além disso, os próprios sistemas de recomendação de produtos podem melhorar a experiência digital dos usuários ou ser usados para fins mais desejáveis, por exemplo, promover mudanças em atitudes ou no autoconceito em direções mais benéficas aos indivíduos (e.g., reduzir o preconceito ou aumentar a autoestima).

É importante ressaltar, ainda, que os resultados encontrados nesta pesquisa apontam para a possibilidade de usar modelos computacionais da personalidade em contextos em que as pessoas estão altamente motivadas a construir uma imagem favorável de si mesmas. Nos contextos citados acima, por exemplo, de perpetuação de *cyberbullying* ou de processos seletivos para vagas empregos, dificilmente as pessoas estão dispostas a revelar aspectos negativos de si mesmas. As medidas indiretas de personalidade podem, nesse sentido, contribuir para capturar respostas mais espontâneas, isto é, mais próximas daquilo que uma pessoa realmente é (e.g., Carvalho et al., 2022). Nesse sentido, modelos de personalidade baseados em rastros digitais podem se mostrar úteis para acessar o fator neuroticismo ou os traços que compõem a *Dark Triad*, por exemplo, por serem consideradas características socialmente indesejáveis.

Apesar de ter-se proposto a realizar uma revisão de literatura mais abrangente do que as que estão atualmente disponíveis para o Brasil (e.g., Limas et al., 2014), algumas limitações podem ser destacadas. Por exemplo, a busca em bases de dados foi restrita a estudos que acessaram rastros digitais por meio de procedimentos automáticos, isto é, diretamente das redes sociais on-line. Esse critério de inclusão foi estabelecido tendo em vista o interesse em encontrar estudos que usaram a observação naturalística em ambientes digitais como instrumento de medida. Sabe-se, entretanto, que o autorrelato também pode ser usado para capturar comportamentos on-line. Em uma amostra brasileira, por exemplo, a intensidade de uso do *Instagram* foi acessada por meio de um instrumento de autorrelato e se correlacionou positivamente com o fator neuroticismo (Fagundes et al., 2021). Esse instrumento visa a questionar as pessoas sobre o quão intensamente elas usam o Instagram, diferentemente dos estudos encontrados nesta pesquisa que acessaram comportamentos on-line diretamente das redes sociais on-line.

É possível, nesse sentido, que esta pesquisa tenha deixado de capturar outros estudos empíricos nacionais e internacionais que observaram a manifestação da personalidade em ambientes digitais. Isso porque alguns estudos recorreram à percepção das pessoas sobre os próprios comportamentos on-line, em vez dos comportamentos em si (*real-life behaviors*) (e.g., Fagundes et al., 2021; Sanecka, 2017). Mesmo que esses estudos não tenham acesso direto aos comportamentos, estando mais suscetíveis a falseamentos e outros vieses, eles fornecem *insights* sobre como a personalidade se manifesta em ambientes digitais. Desse modo, encoraja-se que estudos futuros de revisão abarquem artigos que usaram outros métodos além da observação naturalística para acessar comportamentos on-line. Ainda, estudos futuros podem abarcar um recorte temporal mais recente e investigar redes sociais como *Tik Tok*, que por ser mais recente não foi incluída nesta revisão, e *Instagram*, que não apareceu nos resultados dos artigos encontrados.

Por fim, destaca-se a importância de os psicólogos brasileiros, especialmente os pesquisadores, acompanharem as discussões e os avanços internacionais na área de observação naturalística de características psicológicas em ambientes digitais. O uso de *Big Data* em pesquisas comportamentais tem sido útil para a automação de processos que podem beneficiar a área da psicologia (e.g., Balakrishnan et al., 2019; Faliagka et al., 2014). Além dos já citados, ressalta-se a automação de coleta de dados em pesquisas, por exemplo, pois os participantes podem economizar tempo autorizando a liberação de rastros digitais, em vez de responderem a baterias longas de testes. Como toda tecnologia, entretanto, o *Big Data* pode ser usado para diferentes fins, inclusive de modo a invadir a privacidade de indivíduos (Grassegger & Krogerus, 2017). Quanto mais os psicólogos se envolverem nessa área, desenvolvendo procedimentos técnicos e éticos mais claros, maior a probabilidade desse método de avaliação ser usado para promover o bem-estar e a saúde psicológica.

Considerações Finais

De maneira geral, esta pesquisa mostrou que os rastros digitais se correlacionam de maneira significativa com traços de personalidade e, em alguns estudos, têm sido usados para acessar a personalidade sem recorrer a baterias tradicionais de testes psicológicos. Os resultados encontrados evidenciaram que as aplicações dessa prática são variadas, desde sistemas automatizados para triagem de candidatos a vagas de emprego até sistemas de detecção de *cyberbullying*. Ao mesmo tempo que esses sistemas podem facilitar e aprimorar o processo de avaliação psicológica, ainda é necessário discutir questões éticas relacionadas ao uso de rastros digitais como amostras do comportamento. Por se tratar de uma área nova e em constante atualização, em Psicologia, são desejáveis outros estudos de revisão que incluam redes sociais mais recentes (e.g., *Tik Tok*) e possíveis novas formas de aplicação do uso de rastros digitais na avaliação psicológica.

Referências

- Allport, F. H., & Allport, G. W. (1921). Personality traits: Their classification and measurement. *Journal of Abnormal Psychology and Social Psychology*, *16*(1), 6-40. <https://doi.org/10.1037/h0069790>
- Allport, G. W. (1966). *Personalidade: Padrões e desenvolvimento*. Herder.
- Asadzadeh, L., & Rahimi, S. (2017). Analyzing Facebook activities for personality recognition. *Proceedings of 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 960-964. <https://doi.org/10.1109/ICMLA.2017.00-29>
- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380-393. <https://doi.org/10.1037//0022-3514.83.2.380>
- Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, *124*, 150-159. <https://doi.org/10.1016/j.paid.2017.12.018>
- Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012). Personality and patterns of Facebook usage [Paper presentation]. *Proceedings of the 3rd Annual ACM Web Science Conference* (pp. 24-32). New York.
- Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using Twitter users psychological features and machine learning. *Computers & Security*, *90*. <https://doi.org/10.1016/j.cose.2019.101710>
- Balakrishnan, V., Khan, S., Fernandez, T., & Arabnia, H. R. (2019). Cyberbullying detection on Twitter using Big Five and Dark Triad features. *Personality and Individual Differences*, *141*, 252-257. <https://doi.org/10.1016/j.paid.2019.01.024>
- Barenbaum, N. B., & Winter, D. G. (2010). History of modern personality theory and research. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114-158). Guilford Press.
- Barker, R. G., & Wright, H. F. (1951). *One boy's day. A specimen record of behavior*. Harper & Row.
- Block, J. H., Fisch, C., Obschonka, M., & Sandner, P. (2019). A personality perspective on business angel syndication. *Journal of Banking and Finance*, *100*, 306-327. <https://doi.org/10.1016/j.jbankfin.2018.10.006>
- Carvalho, N. M., Bastos, R. V. S., & Natividade, J. C. (2022). Evidence of validity and accuracy for an implicit measure to assess depressive traits. *Trends in Psychology*. <https://doi.org/10.1007/s43076-022-00145-y>
- Celli, F., & Rossi, L. (2015). Long chains or stable communities? The role of emotional stability in Twitter conversations. *Computational Intelligence*, *31*, 184-200. <https://doi.org/10.1111/coin.12023>
- Celli, F., & Zaga, C. (2013). Be conscientious, express your sentiment! *Training*, *5747*(495).
- Chen, L., Gong, T., Kosinski, M., Stillwell, D., & Davidson, R. L. (2017). Building a profile of subjective well-being for social media users. *PLoS ONE*, *12*(11). <https://doi.org/10.1371/journal.pone.0187278>
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (2007). *NEO PI-R: Inventário de personalidade NEO revisado e inventário de cinco fatores NEO revisado NEO-FFI-R* [versão curta]. Vetor Editora Psico-Pedagógica.
- Costa, P., Jr., & McCrae, R. (2008). The revised NEO Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Sakelofsky (Eds.), *The SAGE Handbook of Personality Theory and Assessment: Volume 2. Personality Measurement and Testing*, (pp. 179-198). SAGE. <https://doi.org/10.4135/9781849200479.n9>
- Craik, K. H. (2000). The lived day of an individual: A person-environment perspective. In W. B. Walsh, K. H. Craik, & R. H. Price (Eds.), *Person-environment psychology: New directions and perspectives* (pp. 233-266). Erlbaum.
- Eichstaedt, J. C., & Weidman, A. (2020). Tracking fluctuations in psychological states using social media language: A case study of weekly emotion. *European Journal of Personality*, *34*, 845-858. <https://doi.org/10.1002/per.2261>

- Elgendy, N., & Elragal, A. (2014). Big Data Analytics: A literature review paper. *Data Mining Applications and Theoretical Aspects*, 214–227. https://doi.org/10.1007/978-3-319-08976-8_16
- Fagundes, L. S., Marot, T. A., & Natividade, J. C. (2021). Use of Instagram, social comparison, and personality as predictors of self-esteem. *PSICO-USF*, 25(4), 711-724. <https://doi.org/10.1590/1413/82712020250410>
- Faliagka, E., Iliadis, L., Karydis, I., Rigou, M., Sioutas, S., Tsakalidis, A., & Tzimas, G. (2014). On-line consistent ranking on e-recruitment: Seeking the truth behind a well-formed CV. *Artificial Intelligence Review*, 42, 515-528. <https://doi.org/10.1007/s10462-013-9414-y>
- Funk, M. (2016, November 19). Cambridge Analytica and the Secret Agenda of a Facebook Quiz. *The New York Times*. <https://www.nytimes.com/2016/11/20/opinion/cambridge-analytica-facebook-quiz.html>
- Gibney, E. (2018, March 29). The scant science behind Cambridge Analytica's controversial marketing techniques. *Nature*. <https://www.nature.com/articles/d41586-018-03880-4>
- Goldberg, L. R. (1992). The development of markers for the Big-Five Factor Structure. *Psychological Assessment*, 4(1), 26-42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84–96. <https://doi.org/10.1016/j.jrp.2005.08.007>
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, 82, 379-398. <https://doi.org/10.1037/0022-3514.82.3.379>
- Grassegger, H., & Krogerus, M. (2017, January 28). The data that turned the world upside down. *Vice*. https://www.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win
- Guntuku, S. C., Zhou, J. T., Roy, S., Lin, W., & Tsang, I. W. (2018). Who likes what and, why? Insights into modeling Users' personality based on image 'Likes'. *IEEE Transactions on Affective Computing*, 9(1), 130–143. <https://doi.org/10.1109/TAFFC.2016.2581168>
- Hagan, C., Carpenter, J., Ungar, L., & Preotiuc-Pietro, D. (2017). Personality profiles of users sharing animal-related content on social media. *Anthrozoos*, 30, 671–680. <https://doi.org/10.1080/08927936.2017.1370235>
- Hall, A. N., & Matz, S. C. (2020). Targeting item-level nuances leads to small but robust improvements in personality prediction from digital footprints [Special issue]. *European Journal of Personality*, 34(5). <https://doi.org/10.1002/per.2253>
- Hall, M., & Caton, S. (2017). Am I who I say I am? Unobtrusive self-representation and personality recognition on Facebook. *PloS One*, 12(9), e0184417.
- Hern, A. (2018, May 6). Cambridge Analytica: How did it turn clicks into votes? *The Guardian*. <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>
- Hughes, D. J., Rowe, M., Batey, M., & Lee, A. (2012). A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Computers in Human Behavior*, 28, 561-569. <https://doi.org/10.1016/j.chb.2011.11.001>
- Hutz, C. S., Nunes, C. H., Silveira, A. D., Serra, J., Anton, M., & Wieczorek, L. S. (1998). O desenvolvimento de marcadores para a avaliação da personalidade no modelo dos cinco grandes fatores. *Psicologia: Reflexão e Crítica*, 11(2), 395-411. <https://doi.org/10.1590/S0102-79721998000200015>
- International Business Machines Corporation. (2018). *Personality insights - API reference | IBM Watson Developer Cloud*. <https://www.ibm.com/watson/developercloud/personality-insights/api/v3/curl.html>
- John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality - A historical review of trait taxonomic research. *European Journal of Personality*, 2(3), 171-203. <https://doi.org/10.1002/per.2410020302>
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory-Versions 4a and 54*. Institute of Personality and Social Research.
- Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., Kosinski, M., Ramones, S. M., & Seligman,

- M. E. P. (2014). The online social self: An open vocabulary approach to personality. *Assessment*, 21, 158–169. <https://doi.org/10.1177/1073191113514104>
- Kim, Y., & Kim, J. H. (2018). Using computer vision techniques on Instagram to link user's personalities and genders to the features of their photos: An exploratory study. *Information Processing & Management*, 54(6), 1101–1114. <https://doi.org/10.1016/j.ipm.2018.07.005>
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805. <https://doi.org/10.1073/pnas.1218772110>
- Lane, J. E. (2016). Big data and anthropology: Concerns for data collection in a new research context. *JASO*, 3(1), 74–88.
- Lima, A. C. E., & Castro, L. N. (2014). A multi-label, semi-supervised classification approach applied to personality prediction in social media. *Neural Networks*, 58, 122–130. <https://doi.org/10.1016/j.neunet.2014.05.020>
- Limas, A. F., Primi, R., & Carvalho, L. F. (2014). Avaliação da personalidade por redes sociais online: Uso do Facebook na área. *Revista Sul Americana de Psicologia*, 2(1), 1–25.
- Liu, Y., Wang, J., & Jiang, Y. (2016). PT-LDA: A latent variable model to predict personality traits of social network users. *Neurocomputing*, 210, 155–163. <https://doi.org/10.1016/j.neucom.2015.10.144>
- Mairesse, F., & Walker, M. (2007). PERSONAGE: Personality generation for dialogue [Paper presentation]. *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics* (pp. 496–503), Prague, Czech Republic, Association for Computational Linguistics. <https://aclanthology.org/P07-1063>
- Mairesse, F., Walker, M., Mehl, M., & Moore, R. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, 30, 457–500. <https://doi.org/10.1613/jair.2349>
- Marinucci, A., Kraska, J., & Costello, S. (2018). Recreating the relationship between subjective wellbeing and personality using machine learning: An investigation into Facebook online behaviors. *Big Data and Cognitive Computing*, 2(3). <https://doi.org/10.3390/bdcc2030029>
- Marouf, A. A., Hasan, M. K., & Mahmud, H. (2020). Comparative analysis of feature selection algorithms for computational personality prediction from social media. *IEEE Transactions on Computational Social Systems*, 1–13. <https://doi.org/10.1109/TCSS.2020.2966910>
- Matz, S., Kosinski, M., Nave, G., & Stillwell, D. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences (PNAS)*, 114, 12714–12719. <https://doi.org/10.1073/pnas.1710966114>
- McCrae, R. R., & Costa, P. T., Jr. (2010). The Five-Factor Theory of Personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159–181). Guilford Press.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90, 862–877. <https://doi.org/10.1037/0022-3514.90.5.862>
- Mori, K., & Haruno, M. (2020). Differential ability of network and natural language information on social media to predict interpersonal and mental health traits. *Journal of Personality*, 89(4), 1–16. <https://doi.org/10.1111/jopy.12578>
- Natividade, J. C., & Hutz, C. S. (2015). Escala reduzida de descritores dos cinco grandes fatores de personalidade: Prós e contras. *Psico (PUC-RS)*, 46(1), 79–89. <https://doi.org/10.15448/1980-8623.2015.1.16901>
- Natividade, J. C., & Hutz, C. S. (2016). Personal characteristics associated with sexuality can be classified into seven dimensions in Brazil. *Personality and Individual Differences*, 97, 88–97. <https://doi.org/10.1016/j.paid.2016.03.030>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>

- Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. P. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*, 934-952. <https://doi.org/10.1037/pspp0000020>
- Pennebaker, J. W., & Francis, M. E. (1999). *Linguistic Inquiry and Word Count: LIWC*. Lawrence Erlbaum.
- Pereira, R. A., Jr., & Inkpen, D. (2017). Using cognitive computing to get insights on personality traits from Twitter messages [Conference paper]. *Advances in Artificial Intelligence - 30th Canadian Conference on Artificial Intelligence Proceedings*, Edmonton, Canadian. https://doi.org/10.1007/978-3-319-57351-9_32
- Pervin, L. A., & John, O. P. (2004). *Personalidade: Teoria e pesquisa* (8ª ed.). Artmed.
- Primi, R. (2018). Avaliação Psicológica no Século XXI: De onde viemos e para onde vamos. *Psicologia: Ciência e Profissão, 38*, 87-97. <https://doi.org/10.1590/1982-3703000209814>
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203-212.
- Receptiviti Inc. (2017). *Receptiviti API user manual*.
- Sanecka, E. (2017). The dark side of social media: Associations between the Dark Triad of personality, self-disclosure online and selfie-related behaviors. *Journal of Education Culture and Society, 7*(2), 71-88. <https://doi.org/10.15503/jecs20172.71.88>
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open vocabulary approach. *PLOS ONE, 8*, e73791. <https://doi.org/10.1371/journal.pone.0073791>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*(1), 117-143. <https://doi.org/10.1037/pspp0000096>
- Tareaf, R. B., Alhosseini, S. A., & Meinel, C. (2019). Facial-Based Personality Prediction Models for Estimating Individuals Private Traits. *IEEE International Conference on Big Data and Cloud Computing (BdCloud)*, 1586-1594. <https://doi.org/10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00233>
- Tuton, T. L., & Solomon, M. R. (2015). *Social Media Marketing* (2nd ed.). Sage Publication.
- Vazire, S., & Gosling, S. D. (2004). e-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology, 87*, 123-132. <https://doi.org/10.1037/0022-3514.87.1.123>
- Winkler, H. J., Rieger, V., & Engelen A. (2020). Does the CMO's personality matter for web traffic? Evidence from technology-based new ventures. *Journal of the Academy of Marketing Science, 48*(2), 308-330. <https://doi.org/10.1007/s11747-019-00671-9>
- Yang, H. C., & Huang, Z. R. (2019). Mining personality traits from social messages for game recommender systems. *Knowledge-Based Systems, 165*, 157-168. <https://doi.org/10.1016/j.knsys.2018.11.025>
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of the Sciences, 112*, 1036-1040. <https://doi.org/10.1073/pnas.1418680112>

Recebido: 13/12/2022
1ª revisão: 30/06/2023
2ª revisão: 11/08/2023
Aceite final: 18/08/2023

