

Edição v. 38
número 3 / 2019

Contracampo e-ISSN 2238-2577
Niterói (RJ), 38 (3)
dez/2019-mar/2020

A Revista Contracampo é uma revista eletrônica do Programa de Pós-Graduação em Comunicação da Universidade Federal Fluminense e tem como objetivo contribuir para a reflexão crítica em torno do campo midiático, atuando como espaço de circulação da pesquisa e do pensamento acadêmico.

BOTS COMO AGENTES DE EXPRESSÃO: Regime de visibilidades e o poder de criar redes

BOTS AS AGENTS OF EXPRESSION: Regimes of visibility and the power to create networks

LORENA LUCAS REGATTIERI

Universidade Federal do Rio de Janeiro (UFRJ). Rio de Janeiro, Rio de Janeiro, Brasil. E-mail: lorenaufes@gmail.com. ORCID: 0000-0002-8747-1053.

AO CITAR ESTE ARTIGO, UTILIZE A SEGUINTE REFERÊNCIA:

REGATTIERI, Lorena Lucas. BOTS COMO AGENTES DE EXPRESSÃO: Regime de visibilidades e o poder de criar redes. Contracampo, Niterói, v. 38, n.2, p. 130-149, ago./nov. 2019.

Enviado em: 12/4/2019. Revisor A: 2/8/2019; Revisor B: 17/7/2019; Revisor A: 21/8/2019. Aceito em: 22/08/2019.

DOI – <http://dx.doi.org/10.22409/contracampo.v38i3.28504>

Resumo¹

O que aprendemos a respeito da utilização de robôs nas redes sociais e seus efeitos comunicacionais? Considerando as práticas midiáticas contemporâneas, os estudos críticos algorítmicos vêm debatendo uma transformação nos regimes de visibilidade (Magalhães, 2018). Este artigo retoma um estudo de caso das eleições presidenciais de 2014 sobre o uso de bots no Twitter como agentes de expressão. Ao coletar dados digitais do Twitter, partiu-se de uma técnica quali-quantitativa de análise das redes sociais para cartografar as estratégias de computacionais de propaganda. Assim, sob efeito dos *bots*, as modulações produzidas da interação entre atores humanos e não humanos fornecem novos parâmetros para compreender fenômenos políticos-comunicacionais.

Palavras-chave

Bots; Visibilidade; Inteligência artificial; Análise de redes sociais.

Abstract

What have we learned about the use of robots in social networks and their communication effects? Considering the contemporary media practices, critical algorithmic studies have been debating a transformation in the regimes of visibility (Magalhães, 2018). This article takes up a case study of the 2014 presidential elections on the use of bots on Twitter as agents of expression. When collecting digital data from Twitter, we process with a qualitative-quantitative technique of analyzing social networks to map computational propaganda strategies. Thus, under the effect of bots, the modulations produced from the interaction between human and non-human actors provide new parameters for understanding political-communicational phenomena.

Keywords

Bots; Visibility; Artificial intelligence; Social network analysis.

¹ Esse artigo é fruto de atualizações inéditas da dissertação e novas investigações da pesquisa de doutoramento.

Panorama do uso de *bots* no espaço digital

O uso de *bots* como ferramentas - sua fusão desde uma tecnologia de agente em uma nova categoria de “agentes críveis” - requer um olhar mais atento sobre o que está por trás da tendência humana de antropomorfizar. Por um lado, dotar programas de software, caráter e personalidade os torna mais amigáveis. Pensar neles como seres semelhantes esfrega as ásperas bordas da crua realidade da interface homem-computador. [...]. Mas há um porém. O ato de personificação obscurece uma importante verdade sobre o relacionamento humano com as ferramentas de software. A dicotomia criada entre “nós” e “eles” é falsa. Estamos juntos nisso. As ferramentas de software - se as chamamos de *bots* ou agentes, sejam eles personagens críveis ou não - são extensões fora do nosso dia a dia de carne e osso (LEONARD, 1997, p. 104, tradução nossa).

Em seu livro, *Bots: The origin of new species* (1999), Andrew Leonard descreve os *bots* como programas de software autônomos e supostamente inteligentes que fornecem um serviço. Leonard, jornalista e autor de colunas de tecnologia da revista *Wired*, nos coloca em contato com uma narrativa histórica e crítica sobre o impacto desses agentes na sociedade. O livro apresenta questões cruciais para o debate na computação e na comunicação, pois revela o papel do software como serviço (SaaS), dos protocolos de comunicação entre a internet e o software (API) e da evolução do processamento da linguagem natural (NLP). Os chamados *chatbots* se popularizaram em serviços de atendimento ao cliente, processos de recrutamento e em plataformas de redes sociais. ELIZA² é considerado o primeiro *chatbot* na história da ciência da computação. Muito embora, em 1966, o termo *chatbot* não existisse. A definição aparece apenas em 1994 com o cientista da computação Michael Mauldin. A ELIZA operava identificando palavras ou termos-chave da entrada para responder ao usuário usando as palavras ou termos-chave de respostas pré-programadas. Por exemplo, se um humano disser em um tweet ou em uma mensagem “Não consigo utilizar o serviço de transferência de conta corrente do banco” e se o banco citado no conteúdo possuir um serviço de *bots*, as palavras-chave são reconhecidas e processadas para que o robô possa responder com mensagens roteirizadas pelo desenvolvedor. Esse tipo de procedimento criou uma ilusão de compreensão e interação com um ser humano real, embora o processo fosse mecanizado.

Essa história é fundamental para compreender as transformações no mundo dos agentes, seus efeitos na comunicação e na sociedade. Nesse artigo, busco alinhar as ciências da computação com as ciências da comunicação para compreender as operações e as implicações desses serviços de software no regime de visibilidades e subjetivação. Busco analisar como as interações entre pessoas e *bots* modulam a experiência das pessoas e contaminam os ambientes digitais de conversação on-line. Assim, denomino *bots* como agentes de expressão das rotinas computacionais interagindo via protocolos de comunicação nas plataformas de redes sociais com o objetivo de potencializar e modelizar o comportamento das conversas e dos humanos.

A literatura especializada na ciência da computação oferece vasto arcabouço de metodologias para detecção de *spams* e *bots*. Neste artigo investigo os robôs nas redes sociais como ferramentas algorítmicas e distintas abordagens para debater os efeitos desses agentes no espaço comunicacional. O desenho de agentes autônomos como os *bots* pode nos dar pistas para investigar o modelo de propaganda em rede, bem como os disparos em massa em outras redes sociais, como, por exemplo, o WhatsApp. No caso dos *bots*, uma variedade de trabalhos pesquisou em diferentes contextos políticos e sociais o modo como essas operações de informação são planejadas e coordenadas para causar a desorientação da audiência.

² Para saber mais sobre a história do chatbot ELIZA, ver: <https://www.analyticsindiamag.com/story-eliza-first-chatbot-developed-1966/>.

O objetivo dessa pesquisa é analisar certos aspectos do padrão de propagação e ativação de circulação de informação sob efeito dos *bots*, especificamente das *botnets*.

Desde os anos 1990, estudos dedicaram-se à complexidade dos agentes e de mecanismos de seleção decisório dos modelos de inteligência artificial (MAES, 1990), testando e simulando condições de performance e julgamento de ações na medida em que se adaptam a novos ambientes, criando modelos algorítmicos capazes de proceder na tomada de decisões e considerando a escalada de dados e a complexidade de variações das metas ao longo do tempo. Esses trabalhos acabaram oferecendo uma reflexão sobre o lugar da agência em atividades cognitivas e do que se entende por inteligência, analisando funcionalidades emergentes dos agentes autônomos e se a inteligência artificial seria capaz de corporificar suficiente inteligência para expressar, captar, se adaptar e distribuir ações. Essas pesquisas lideradas por Patti Maes revolucionaram a inteligência artificial pela abordagem situada da arquitetura e do desenho de agentes autônomos, tornando-se menos generalista e mais realista.

A onda de estudos da computação se aliou à comunicação, notadamente nas investigações envolvendo o uso da propaganda computacional em cenários políticos pelo mundo (HOWARD, WOOLEY e CALO, 2018; FERRARA, 2017; WOOLEY, KOLLANYI e HOWARD, 2016; HOWARD e KOLLANY, 2016). O relatório do *Oxford Internet Institute, Computational Propaganda in Brazil: Social Bots during Elections* (ARNAUDO, 2017) reuniu trabalhos que focam nos processos computacionais envolvidos na elaboração das contas automáticas e no cálculo da influência desses processos automatizados nas redes sociais, bem como no aperfeiçoamento das ferramentas de detecção de robôs (SAVVOPOULOS et al., 2018; VOSOUGUI, S.; ROY, D.; ARAL, 2018; CHAKRABORTY et al., 2017; SILVA et al., 2016; MESSIAS et al., 2013; BENEVENUTO et al., 2010). No caso das investigações realizadas pelo LABIC (Laboratório de Estudos em Imagem e Cibercultura), Arnaudo (2017) destacou o impacto dos *bots* nas eleições presidenciais de 2014, a votação do projeto de lei do Marco Civil da Internet, o processo de impeachment com o *Vem pra Rua* e as eleições municipais do Rio de Janeiro em 2016 (REGATTIERI, 2016; REGATTIERI et al., 2014; MALINI, 2016; CÔRTEZ et al., 2016).

Esse conjunto de pesquisas evidencia a utilização do *bot* no ecossistema informacional das plataformas mediadas por algoritmos como uma ferramenta algorítmica funcionando em meio a culturas algorítmicas (GILLESPIE, 2017). Desde então, os *bots*, *spams* e os *trolls* no campo da comunicação são estudados para compreender os impactos enquanto instrumento da amplificação da desinformação e conteúdo de baixa-credibilidade, propagação de campanhas de cunho político, social e cultural e da radicalização e polarização do debate público nas redes sociais. O estudo da arte das investigações sobre identificação de *bots* pode ser ancorado na plataforma *Botometer* (VAROL, 2017). A plataforma verifica a atividade de uma conta do Twitter e atribui uma pontuação baseada na probabilidade da conta ser um *bot*. Pontuações mais altas são mais parecidas com *bots*. Portanto, o processo de automação na emulação de perfis nas redes sociais on-line possui um método computacional de detecção, estimativa e caracterização. A plataforma consiste em detectar o que chamam de robôs sociais no Twitter, a partir de recursos extraídos de dados públicos e meta-dados sobre usuários nas redes. São eles: amigos, conteúdo e sentimento de tweets, padrões de rede e séries temporais de atividades. Ao passar pelo crivo da plataforma, a depender da performatividade do perfil, há uma probabilidade de um perfil ser considerado robô. E é verdade também que um perfil humano, ao passar pela avaliação da ferramenta, pode vir a inferir que esse mesmo perfil é um robô. Assim, é uma relação com a probabilidade, desde que certas características tidas como parte dos processos de automação – conforme delineadas pela equação matemática – indique o quanto um usuário humano performa as características de um robô.

Os *bots* são contas controladas por softwares e essas contas performam tarefas automatizadas, podendo postar conteúdo, interagir entre si e com usuários humanos e outros bots, por meio de conexões on-line – assim como pessoas reais. Sanger (2018); Jamieson (2018); Benkler, Farris e Roberts (2018) e Phillips e Milner (2017) debatem como a ação de modelos de propaganda nas redes sociais pode funcionar

como armas de organização, modelização, desinformação e radicalização capazes de interferir na opinião pública e, eventualmente, alterar o curso dos processos eleitorais e dos debates na democracia. Na busca por evidências quantitativas, Shao et al. (2018) analisaram a difusão viral da desinformação a partir de 14 milhões de mensagens que espalharam 400 mil artigos no Twitter, durante dez meses, entre 2016 e 2017, nos Estados Unidos. A metodologia usa dados qualitativos cedidos por agências de checagem de fatos legítimos para classificar parâmetros relacionados aos conteúdos de alta e baixa credibilidade. Desse modo, encontraram evidências de que os *bots* desempenharam um papel desproporcional na divulgação de artigos de fontes de baixa credibilidade. Os resultados indicaram que os *bots* operaram na amplificação desse conteúdo nos primeiros momentos de divulgação, antes que o conteúdo se tornasse viral. Ao mesmo tempo, os *bots* também funcionavam segmentando usuários com muitos seguidores por meio de respostas e menções. Os perfis humanos são vulneráveis a essa manipulação, compartilhando novamente o material postado por *bots* e passando a compor juntos – cada qual incorporando partes de suas conexões em rede. Os *bots* são o suporte dos sites de baixa credibilidade, pois estrategicamente são o oxigênio da amplificação (PHILLIPS, 2018).

Como em uma política armamentista na internet – uma *cyberwar* adotando e adaptando modelos de propaganda em rede com fins específicos –, a automação para a modelização do debate público nas redes sociais on-line utilizam estratégias coordenadas usando os chamados *bots* para formar manadas de opinião. Tão importante quanto identificar os perfis robôs controlados por software é entender o impacto dessa guerra no espaço comunicacional e como essas ferramentas transformam o regime de visibilidade (MAGALHÃES, 2018); as contaminações nos processos de subjetivação e na formação de novos modelos de conversação e propaganda (SANGER, 2018; BENKLER, FARIS e ROBERTS, 2018). E se alguns o fazem com bancos de dados privilegiados? Ou se há financiamento para investir em ferramentas de inteligência artificial e aprendizado da máquina? O que seria da opinião pública diante de conteúdos maliciosos, teorias conspiratórias, boatos, notícias falsas ou enganosas? (MARWICK e LEWIS, 2018). E, ainda, o que modula o sujeito alvo dessas estratégias de propaganda? Nesses processos emergem os perfis ciborgues. Eles podem funcionar como *uma máquina de influência digital (DIM)* (NADLER, CRAIN e DONOVAM, 2018), viabilizando a gestão e a manutenção de conteúdos criados para propagar mensagens nas redes sociais, caracterizando-se por uma “infraestrutura de coleta de dados e capacidades de segmentação” (NADLER, CRAIN e DONOVAM, 2018), desenvolvida por plataformas de anúncios, editores da web e outros intermediários. O funcionamento dessa máquina inclui um esquema montado para o monitoramento do usuário, direcionamento do público-alvo e utilização de tecnologias automatizadas que aumentam seu alcance e, em última análise, seu poder de influência. Segundo os autores, três mudanças fundamentais no panorama midiático dos Estados Unidos proporcionaram as condições para a consolidação desse modelo de propaganda: o declínio do jornalismo profissional, a expansão dos recursos financeiros dedicados à influência política e a crescente sofisticação da publicidade direcionada com pouca supervisão.

Pode-se inferir que as *botnets* operam uma intervenção precoce que expõe muitos usuários a informações de baixa credibilidade, aumentando as chances de uma mensagem ser “viral”. A consolidação da amplificação do conteúdo e o sucesso do modelo de propaganda são marcados pela entrada dos perfis humanos na rede de difusão. Essa dinâmica importa em seus efeitos de expressão, dentro do espaço de comunicação do Twitter e da propaganda em rede. Reviso uma série de estudos entre a comunicação e a computação para compreender como os dados digitais extraídos das redes sociais on-line podem indicar evidências de organização, distribuição e modelização de grupos nas redes sociais.

Na próxima parte do texto, analiso a relação entre a simulação de realidade virtual e os modos de expressão do sujeito no espaço digital em busca de visibilidade, tratando a virtualidade como terreno de elaboração ética. A construção da tomada de decisões por humanos no espaço virtual considera as condicionantes dessas ferramentas algorítmicas levando em consideração o ambiente em que estão inseridos. A partir de dados digitais das eleições presidenciais de 2014, analiso características de

comportamento dos *bots* que colocam em ação tarefas coordenadas para atingir objetivos específicos, considerando as dinâmicas de interação do próprio ambiente. O processo de execução e performance de ações específicas, planejamento e deliberação joga luz sobre a arquitetura e o desenho de *bots* enquanto agentes de expressão. Por fim, esse trabalho ainda em andamento, dispõe de visualizações em grafo para mostrar que *botnets* estabelecem novas dinâmicas de comunicação entre humanos e máquinas.

Ferramentas algorítmicas de propaganda em rede: pistas sobre a circulação da informação, subjetivação e mediações de visibilidade

Bots, como todas as criaturas, pertencem ao seu ambiente. E, neste caso, o ambiente é a Net. Os MUDs (domínios de multiusuários) que se proliferam, os milhares de quadros de avisos que compõem as notícias da Usenet, a praticidade das infinitas salas de chat da rede IRC, a inexorável circulação e transformação da World Wide Web – apenas na mais acidental das maneiras se pode dizer que a Net possa ter sido planejada ou organizada ou arregimentada. É um exemplo de grande sucesso do poder de distribuição paralela – infinitamente criativo, infinitamente mudando. [...] A multiplicidade de ambientes de rede é uma das razões por que a tentativa de discernir uma genealogia particular ou anatomia de *bots* é uma missão cheia de perigos de ordem semântica e logística. *Bots* não compreendem apenas uma nova espécie, mas um espectro completo de novas espécies, um novo filo sob o sol digital (LEONARD, 1997, p. 17, tradução nossa).

No plano do social, longe de uma fascinação pela técnica ou de reforçar um horizonte catastrófico em que as máquinas nos levam a uma situação de inumanidade e perda das dimensões éticas da vida, proponho territorializar essa relação maquinica entre humanos e não-humanos. Guattari (2015) investe numa tentativa de reconceitualizar a máquina, posicionando-a num cruzamento entre a inércia e a invenção; por um lado, o caráter vazio, por outro, como um objeto de individuação subjetiva e subjetividade coletiva. Nessa humanização robótica ou a robótica humanização, os *bots*, agentes autônomos, os vírus, os *bots* sociais, os perfis falsos, as sátiras, desde a perspectiva dos humanos, são vistos como possuidores de uma alma e, às vezes, até de um certo poder de organizar o espaço informacional. Busca-se demonstrar com essa investigação que a retroalimentação entre níveis de informação de natureza afetiva, cognitiva e informacional é o insumo usado pelo robô na produção de uma protosubjetividade (GUATTARI, 2015), que objetiva prolongar a interação entre humano e máquina e estabelecer níveis de visibilidades.

[...] o que proponho não é exatamente um retorno a uma concepção animista, mas, pelo menos, tentar considerar que, na máquina, a interface maquinica, há algo que seria não da ordem do animal, alma ou animal- humana, mas da ordem de uma protosubjetividade. Ou seja, na função de máquina de uma consistência, relacionamento consigo mesmo e relação com a alteridade.⁵⁰ (GUATTARI, 2015, p. 89).

Magalhães (2018) considera o problema da mediação da visibilidade. O desenho das redes diante dos aspectos de simulação dos *bots* nas redes sociais pode gerir ordenamentos difusos da produção de subjetividade. O autor argumenta que “o significado normativo que os usuários finais atribuem às decisões algorítmicas tem propriedades éticas constitutivas” (MAGALHÃES, 2018, p. 8). O ordenamento organizado pela ação coordenada de uma rede de *bots* deve supor a realidade e as percepções que constituem as expressões de um perfil nas redes sociais. Ainda assim, a tentativa de alteração de comportamento e valoração de um modo de ser ético não funda a identidade de um usuário final. A lógica objetiva e a consequência da ação de um perfil não podem ser explicadas pelas normatividades embutidas em um algoritmo, ou até mesmo em um *bot*. As deliberações do sujeito nas redes sociais podem se distinguir nos aspectos da governança algorítmica em um regime de visibilidades, entretanto a avaliação ética da virtude e os aspectos morais podem prejudicar a autonomia do próprio sujeito. Se deliberadamente o

sujeito em contato com o algoritmo está em alteração, as metas, parâmetros e percepções operam novas práticas de visibilidade. Como ferramentas algorítmicas, os *bots* se realizam em suas *botnets* e necessitam da arquitetura de interação humana para circular e orientar novas práticas de visibilidade. Entretanto, *botnets* não regulam éticas, tampouco governam a moral ou estabelecem códigos de conduta.

O contexto central compreende as decisões que os algoritmos tomam em nome dos usuários finais (por exemplo, o feed de notícias do Facebook). Da perspectiva não especializada do usuário final, eles são uma forma de contexto sócio-materializado instanciado não por meio de complexos de códigos computacionais e dados digitais, mas de uma interface enganosamente simples. É a tentativa deliberada de um indivíduo de mudar a si mesmo em resposta ao que ele especula como a lógica por trás dessas decisões que define uma subjetivação ética como “algorítmica” (MAGALHÃES, 2018, p. 7).

Uma das formas mais populares de socialização entre humanos-humanos e humanos-máquinas em um ambiente digital são os jogos de realidade virtual. Pelo menos desde o fim dos anos 1970, ambientes virtuais em rede emulam características do mundo físico e social, permitindo a interação entre usuários em um mundo virtual. O primeiro jogo de emulação virtual foi o Dungeons and Dragons. O Multi Usuário Dungeons ou Multi Usuário Dimensional (MUD) é um software que funciona como uma plataforma virtual que conecta usuários e permite apenas comandos textuais. Os jogadores conectam-se na plataforma e podem escolher o nome, o gênero e também descrever uma pequena biografia. Segundo Curtis (1992), esse ambiente multiparticipante, de realidade virtual extensível e acessível pelos usuários em rede, é um local de encontro virtual com aspectos e hábitos da vida real, como a colaboração entre pessoas para constituir um projeto coletivo. Apesar de a construção artificial de um ambiente e a interação serem basicamente por meio de textos, as sensações têm implicações na vida dos usuários e são entendidas como um fenômeno social de comportamento. Essa experimentação de uma realidade virtual dialoga com o debate atual sobre o tempo que os usuários passam socializando nas redes sociais e a performatividade nessa outra camada social da vida, não tão diferente dos jogadores – para conectar-se ao Twitter, por exemplo, o usuário deve escolher um nome, um avatar e preencher uma biografia.

Os jogadores MUD normalmente gastam seu tempo conectado socializando uns com os outros, explorando as várias salas e outros objetos no banco de dados e adicionando tais objetos desenhados por eles mesmos. Eles variam muito na quantidade de tempo que passam conectados em cada visita, variando de apenas um minuto a várias horas; alguns jogadores ficam ligados (e quase sempre ociosos) por vários dias, apenas ocasionalmente participando ativamente (CURTIS, 1992, p. 4).

Para Turkle (1994), no contexto do habitat virtual de múltiplas dimensões, o MUD, as entidades de inteligência computacional, provoca questões sobre esse status em uma dimensão social. Essas entidades, que optei por caracterizar como bots e não como inteligência artificial, como a autora o faz, podem ser criadas por jogadores corporativos. Esses bots podem ficar vagando pelo MUD e participar do ambiente virtual de modo a induzir os outros jogadores, pois o bot pode fornecer informação, contar piadas, ser grosseiro, dar referências sobre locais equivocados. Quando menos se espera, o jogador pode estar empenhado em uma longa conversa com um bot sem ao menos perceber que todo aquele encontro se deu com uma entidade programada por outro jogador, podendo até ter sido criada para aparecer e persuadi-lo a seguir outro caminho. Resumindo, é como se estivessem participando de um Teste de Turing³ sem terem

³ O Teste de Turing operacionaliza um teste no qual um interrogador (humano) faz perguntas a duas entidades ocultas. Uma das entidades é humana e a outra é uma máquina (computador). O interrogador e as entidades se comunicam indiretamente usando um teclado. Essa operação funciona de modo que o interrogador é testado por meio da condução de um diálogo. Por fim, o interrogador deve decidir qual é o humano. O computador é treinado via programação para imitar um humano, usando de um sistema lógico de alternativas. Se o interrogador não conseguir afirmar qual das entidades é humana, presume-se que a máquina atingiu um certo grau de intelectualidade.

sido avisados. No contexto da realidade virtual dos MUDs, a proposta de Turkle é se os bots deveriam ou não “anunciar” sua artificialidade. No contexto das redes sociais, os perfis fakes que performam um outro “eu” passam por situação semelhante ao interagir com outros usuários. Percebe-se que, em usuários famosos do Twitter, como @Samara7Days, @HugoGloss e @DilmaBolada aparecem facetas identitárias que, desenvolvidas ao longo do tempo, dão “a cara” desses avatares. Visto que imitam alguém ou criam sua própria narrativa inventando um personagem, tanto se aproveitam dessa popularidade, que só é possível por conta da comunidade que se identifica com o perfil, como tem na vida virtual grande parte da vida real. A visibilidade é vital para o prolongamento da vida desses perfis.

As redes sociais *online* como espaço de simulação dessa *protosubjetividade* e também por serem uma plataforma ampla de discussão pública interativa possibilitam todo tipo de ruído. A dimensão criativa do criador e as técnicas de web semântica tornam cada vez mais difícil discernir, nas redes sociais, se quem responde ou retweeta uma mensagem é um *bot* ou não. O Twitter e o Facebook dividem essa peculiaridade, pois são, fundamentalmente, redes que valorizam a popularidade e a influência, o que pode tornar a população usuária dessas redes mais suscetível ao ataque de *bots*. A mecânica das relações com o *bot* é definida pela repetição, enquanto a humana tem a sua centralidade na variação. As percepções (impressões e ideias) a partir de sensações como agradabilidade, empatia e prazer passam pelas paixões, portanto, não está no domínio da racionalidade. Desse modo, a forma pela qual uma pessoa faz distinções entre aquilo que ofende ou não outra pessoa ou um conjunto de pessoas não é a mesma que os robôs. O caso do *bot* Tay nos direciona justamente para essa questão nas relações comunicacionais. Principalmente no caso político, já que o estudo de caso utilizado foram as eleições presidenciais, o engajamento entre os eleitores de Dilma e os de Aécio implica sentimentos completamente diferentes entre as duas comunidades. De que modo um *bot* poderia interagir com outro perfil para ganhar votos? Isso depende da capacidade de se detectar ironia, humor, estranheza e outros tipos de interações. Isso não significa apenas classificar frases como positivas, negativas ou neutras (PATODKAR, V., 2016; CANUTO, GONÇALVES, BENEVENUTO, 2016), mas demanda uma habilidade fora da alçada da lógica computacional. Na verdade, atualmente, este é o grande desafio da inteligência artificial e dos métodos de análise de redes sociais: remontar, na máquina, a base para se fazer distinções relacionadas aos sentimentos.

Ferrara et al. (2016) debatem as características do que chamam de bots sociais, relacionando-os com o contexto, a rede, o sentimento e os padrões temporais de atividade. Afinal, que rastros e marcas, ou seja, qual assinatura os bots deixam nessa engenharia social? O grande desafio é distinguir humano de bot, tendo em vista a mimetização dos robôs sociais que se proliferam nas redes sociais. Segundo os autores, os bots sociais tentam imitar os usuários humanos alcançando uma rede de informações e mídia para preencher seus perfis e postar material previamente coletado em determinados momentos, em geral, momentos em que um assunto específico está nos trending topics. Numa tentativa de emular a temporalidade humana, postam conteúdos em um intervalo de tempo que não cause alerta ao modelo de detecção do algoritmo do Twitter – essa estratégia é uma marca de curadoria e de produção de conteúdo típica dos humanos. Aqui, aparecem duas vertentes de análise: uma se empenha em investigar métodos para detectar os bots; a outra investe no desenvolvimento de técnicas para identificar a comunidade de usuários-alvo dos bots. O resultado avaliou que bots retweetam mais que os humanos e são pouco retweetados. Ao mesmo tempo, identificaram que os bots pouco interagem, pois produzem poucos tweets, respostas e menções. Os bots têm nome de usuário geralmente maior e a data de criação do perfil tende a ser recente.

Wagner et al. (2012) partem de uma posição não muito comum entre os pesquisadores, isto é: entender o comportamento dos usuários nas redes sociais que são atacados por bots para saber que grupos são mais suscetíveis a uma interação automatizada por uma das partes. Para tanto, os autores justificam que é necessário prever a interação de usuários na rede para entender o comportamento e a maneira como uma rede de usuários se conecta, bem como desvendar de que formas os usuários

mantêm relações duradouras nas redes sociais (CHENG et al., 2011; ESTRADA & MENDONZA, 2014). A construção de modelos para identificar usuários mais propensos e prever níveis de suscetibilidade ao ataque dos bots passa pela variável reciprocidade. Diante do comportamento dos usuários enquanto se mantêm conectados na rede, quando os bots “atacam”, exploram três grupos de afinidade, divididos em rede, comportamento e linguística. De modo geral, os resultados apontaram que os usuários ativos em redes sociais mais suscetíveis a influências de bots são aqueles mais comunicativos, engajados em conversas constantemente e que, em suas interações, usam palavras mais afetuosas.

Por outro lado, Freitas et al. (2015) partem do princípio de que o Twitter é uma rede social vulnerável ao ataque dos bots. O Twitter é uma rede popular para a distribuição de conteúdo, principalmente a circulação de URLs. Estudos relacionados ao clique de spam (GRIER et al., 2010) contribuem para detectar a presença de links que direcionem para sites maliciosos. Para verificar em que medida os algoritmos implementados pelo Twitter defendem a plataforma de ataque de bots, os pesquisadores testaram 120 bots com atributos ligados ao gênero da bio do perfil do bot, o nível de postagens, o aplicativo usado para postagens e a comunidade de relação do bot. Os autores confirmaram a hipótese de que os bots conseguem passar pelos métodos do Twitter para detectar usuários falsos de modo relativamente fácil, mesmo aqueles bots que utilizam mecanismos simples de infiltração na rede.

Outros resultados apontaram que o gênero da bio não importa nesses algoritmos de diferenciação entre humano e máquina. A frequência de postagens mostrou que quanto mais ativo um bot, seu grau de semelhança com um comportamento considerado humano pelo método do Twitter é mais alto. No sentido da interação, a questão linguística é quantificável, mesmo que a linguagem dessa plataforma seja majoritariamente informal e com poucos caracteres. Resultados mostraram que usuários humanos no Twitter não foram capazes de distinguir tweets gerados pelos bots. Outro fator importante identificado pelo artigo evidenciou que bots têm alto grau de infiltração em grupos específicos quando são criados com a finalidade de entrar em uma comunidade de usuários bem conectados. Em suma, essa análise nos é cara, pois constrói um bom panorama da pesquisa sobre bots e política: o fato de que o gênero na bio do usuário não adquire tanta relevância nessa agência entre humano e máquina, mas que outras características que mimetizam uma performance humana possuem alto grau de influência, tais como a frequência de postagens, a densidade de conexão em grupos de usuários e a popularidade do bot.

Levando em consideração a importância do Twitter e o volume de dados que circula nessa rede e que serve para o desenvolvimento de outras plataformas de monitoramento – não só política, mas de trânsito, epidêmica, entre outras –, Freitas et al. (2014) desenvolveram um método para detectar os bots e compreender como eles se comportam no Twitter, a partir de atributos linguísticos, de conteúdo e de informação do usuário. A análise apontou que os bots tendem a postar mais tweets contendo URLs e *hashtags* do que fazem os humanos. No campo linguístico, os bots acabam tendo uma prática padronizada, o que resulta em um modelo passível de identificação mais detectável. Sobre o comportamento desses agentes de expressão, os bots em redes sociais não são tão “sociais” assim e, no Twitter, não se engajam em conversas tanto quanto os humanos, ou seja, os bots não produzem tantas respostas ou menções (os “replies”). O trabalho de Miranda Filho (2014) montou uma estrutura metodológica para pesquisas de opinião por meio de dados coletados em redes sociais. Para realizar a tarefa de traduzir a opinião expressa nas redes sociais, desenvolve diferentes etapas em um processo que inclui a caracterização do sexo, classe social, idade, análise de sentimentos e identificação de usuários spammers (nossos conhecidos bots) e jornalísticos integrados à ferramenta que realiza a tarefa de classificação. Depois da coleta, os dados são pré-processados para remoção de acentos, pontuações e stopwords (artigos, preposições e advérbios) e preparados para seguir o esquema seguinte de análise de opinião em redes sociais: primeiro, a informação passa por uma etapa de caracterização da mensagem, definida como análise de sentimentos (o que pensam? Qual a opinião?); o segundo passo é a etapa de caracterização dos usuários, definida como caracterização comportamental (posso confiar? O perfil é verdadeiro? É uma propaganda ou uma notícia?)

e caracterização pessoal (onde está? Qual sexo? Qual idade? Qual classe social?); e terceiro, a etapa de formação da amostra, compondo os dados e a possível margem de erro, apontando para a opinião dos usuários (o que pensam? Quem ganha? Qual a preferência?).

Todavia, essas pesquisas não analisam as transformações das práticas comunicacionais e os seus efeitos nos modos de subjetivação contemporâneo. Os métodos computacionais colaboram com o mapeamento quantitativo e indicam aspectos da qualidade das relações que se modelam no espaço comunicacional. Focam nos resultados de otimização sobre o aprimoramento dos métodos e das características de infiltração dos bots na camada social que se retroalimenta via ambiente digital. A transformação no regime de visibilidade por uma “subjetivação ética algorítmica” (MAGALHÃES, 2018) indica o que pode haver de humanidade na normatividade dos dados digitais. É nesse sentido que esse estudo de caso ressalta a potencialidade da comunicação em produzir novos conhecimentos sobre o tema da interação humano-máquina e seus efeitos no ecossistema informacional, principalmente no tratamento de grande quantidade de dados (SHOREY, HOWARD, 2016).

Metodologia de identificação de *botnets*: Mapeando os agentes de expressão

Mais do que identificar se um perfil é uma entidade autômata, *bot* ou um robô social, importa o modo como esse perfil se agencia em rede alterando a si próprio e o comportamento dos demais perfis, transformando as políticas de circulação das informações em uma rede social. Na pesquisa metodológica, os dados são apresentados como uma série de mapas ou grafos de rede que descrevem o modelo de amplificação de conteúdo durante as eleições presidenciais de 2014, com foco nas candidaturas psdbista e petista, suas relações entre si por diferentes medidas estatísticas e a análise de *botnets* (redes de robôs). Na metodologia apresentada, os mapas contêm nós que representam os perfis dos usuários no Twitter; vértices e links entre os nós que qualificam a relação entre os nós como retweets, menções ou respostas; e, por fim, o mapa é particionado em modulações de cores distintas para descrever a espacialização das comunidades que formam a rede.

O tamanho de cada nó indica sua proeminência em relação com outros nós. No mapa do Twitter, a arquitetura é definida pelas conexões entre os usuários e, nesse caso, prioriza-se o número de vezes que um perfil foi replicado (grau de entrada) e o número de vezes que um perfil replicou outro perfil (grau de saída). Os vértices, ou links entre os nós, formam a arquitetura da rede. Essas conexões representam as relações entre os nós. Para qualquer par de usuários em relação no banco de dados das eleições presidenciais de 2014, quanto maior o número de conexões entre um usuário e outro, mais próximos eles são reunidos pelo algoritmo de modulação. A estrutura geral dos mapas do Twitter é determinada pelos padrões de compartilhamento dos usuários dessa rede social em relação com as candidaturas psdbista e petista.

Destacamos aspectos do algoritmo de modularidade, pois a formação de comunidades no Twitter possibilita compreender a vitalidade presente na circulação de um conteúdo. A modularidade é uma das possíveis medidas para a detecção de comunidades em redes complexas. O objetivo do algoritmo de modularidade (BLONDEL et al., 2008; FORTUNATO, 2010) na detecção de comunidade é identificar o particionamento da rede em comunidades, para que a modularidade na rede seja otimizada. As *botnets* operam nos modelos de propaganda para ampliar a disseminação de conteúdo, expondo a rede a todo tipo de conteúdo com o objetivo de induzir o compartilhamento. Para entender o papel crítico dos *bots* na disseminação viral de conteúdo, Shao et al. (2018) demonstram quantitativamente a eficácia de ataques de disseminação de conteúdos maliciosos. Segundo os pesquisadores, os *bots* são superespalhadores de conteúdo de baixa credibilidade (como notícias falsas, teorias de conspiração e boatos) e postam automaticamente links para artigos, retweetam outras contas ou executam tarefas autônomas mais

sofisticadas, como seguir e responder a outros usuários. Entretanto, o padrão de amplificação dos robôs sugere uma tendência em se envolver em momentos específicos na disseminação de artigos populares. Ou seja, os *bots* são coordenados para uma utilização estratégica e não necessariamente compõe a conversação em todas as etapas de um modelo de propaganda. Os resultados indicam que os *bots* são mais prevalentes nos primeiros segundos após um artigo ser publicado pela primeira vez no Twitter do que em fases posteriores.

Entre os dias 06 de julho e 28 de outubro de 2014, o Laboratório de Estudos sobre Imagens e Cibercultura, da Universidade Federal do Espírito Santo (Labic/Ufes), extraiu tweets com mensagens e imagens relacionadas à eleição para presidência do Brasil, publicadas no site de rede social Twitter. Durante esses meses, período permitido pelo Tribunal Superior Eleitoral para a propaganda dos candidatos, foi coletado cerca de 12 milhões de tweets. Os filtros no trabalho com grande volume de dados são essenciais. Nessa investigação utilizamos, por exemplo: delimitação de uma faixa de tempo, número de tweets únicos, frequência de tweets ao dia, número de retweets, número de respostas, número de menções, total de usuários únicos e total de tweets originais. Desses filtros, podemos ainda observar as *hashtags* indexadas, as *hashtags* únicas, as mensagens que contêm URLs, além de imagens e vídeos. Os termos e *hashtags* coletados nesse processo foram #debatenaaglobo, #13brasiltocomdilha, #Aecio45pelobrasil, #melhorcomdilha, #aeciopelamudança, #somostodosdilha, #votoaeciopelobr45il, #aecio45confirma, #13brasiltodocomdilha, #aecio45pelobrasil, #agoraaecio45confirma, #desesperodaveja, #ripbrasil e #impeachment.

Tabela 1 – Com a linha do tempo de informações sobre o banco de dados coletados a partir da API Search do Twitter

Data	Usuários	Tweets	Original	Retweets	Respostas	Menções	Hashtags
19/10/2014	708	940	278	588	74	206	814
20/10/2014	1992	3945	1553	2149	243	1267	3484
21/10/2014	1255	2404	952	1284	168	671	2177
22/10/2014	12316	93906	18319	71352	4235	27471	81697
23/10/2014	30088	220726	44866	164822	11038	66028	184285
24/10/2014	208825	1237126	391841	786204	59081	243831	694344
25/10/2014	69183	295083	73089	207272	14722	73919	163493
26/10/2014	152146	403232	120409	262823	20000	83275	189857
27/10/2014	44789	90420	29498	56249	4673	18535	32228

Fonte: Twitter

Nos oito dias de coleta pela Search API do Twitter, temos um total de 2.347,782 tweets, resultando numa frequência de 293.473,17 tweets ao dia. Dessa base de dados, apenas 41,523 são “geotageados” (um número considerado até alto nas pesquisas com dados coletados do Twitter). Desses tweets, o número de usuários únicos foi de 348,680. O número de tweets originais é de 680,805, advindos de 196,996 usuários. Desse banco de dados da última semana das eleições, identificamos que 371,645 usuários enviaram e receberam mensagens. Os tipos de tweets são divididos em retweets com 1.552,743 (66,14%) do *dataset* (226,122 remetentes e 59,915 destinatários); tweets com 680,805 (29%) e respostas com 114,234 (4.87%) do *dataset* (38,029 remetentes e 32,839 destinatários). O número de menções foi de 515,203 (97,581 remetentes e 26,859 destinatários). Este último é identificado quando uma menção a um usuário encontra-se no meio do texto do tweet, diferente da resposta ao tweet, quando a menção acontece no começo. As interações durante o período de oito dias somam os ATs, MTs e RTs em 2.182,180 tweets (249,899 remetentes e 91,664 destinatários).

A tradução dessas informações de modo a gerar um material empírico, pronto para ser processado, acontece em colaboração com métodos computacionais de coleta e mineração de dados. A coleta dos *tweets* publicados nesse período foi feita a partir de um script chamado FORD. Essa ferramenta busca os *tweets* que contêm termos previamente selecionados e os armazena em um banco de dados do MongoDB, instalado em um servidor remoto. O próximo passo é o processamento desses *tweets* através do 'FORD parse', armazenados em arquivos separado por vírgula (.csv), capaz de separar esse arquivo em outros três arquivos: uma AT (quando *tweets* iniciam com o @arroba do perfil), MT (menções ao @perfil no meio de texto do *tweet*) e os RTs (retweets como um compartilhamento automático das mensagens).

Para iniciar a extração das perspectivas (MALINI, 2016) contagiadas com os bots, o processo de análise de redes sociais desenvolve com método perspectivista a exploração do arquivo de RTs. O Gephi (BASTIAN, HEYMANN e JACOMY, 2009) é a ferramenta utilizada nas pesquisas do laboratório com extração de grande volume de dados, pois é um software completo para analisar, gerar estatísticas e visualizações. Com o Gephi, produzi relatórios estatísticos e visualizações em uma imagem com o design de um grafo. É a partir de estatísticas e filtros que o que passa pela rede se revela. Essa ferramenta livre pode gerar estatísticas importantes, como o grau de entrada ou de saída, hubs e autoridades, modularidade, centralidade de intermediação, aproximação e autovetor. Esses resultados permitem navegar pelos rastros deixados pelos bots e pelo arcabouço semântico de uma rede.

O segundo passo é aplicar o algoritmo de modularidade para encontrar as perspectivas visualizadas na área central do software, como frações da rede em cores distintas, uma vez que usamos Gephi, baseado no método de Louvain (BLONDEL et al., 2008). Esse método realiza a detecção de comunidades em grafos ponderados e tem características com uma heurística de captura acelerada, com a otimização do local de modularidade. O método de Louvain é um algoritmo que encontra partições de alta modularidade em grandes redes em curto espaço de tempo e faz passes que consistem em duas etapas: primeiro, a otimização de modularidade local, ordem aleatória de vértices varrendo toda a rede; segundo, a agregação da comunidade onde os passes se repetem de forma iterativa, até que a modularidade é maximizada e nenhum aumento é possível. Visualmente os clusters ou sub-grafos indicam o contexto de formação dessa aliança entre os perfis – sejam humanos ou não-humanos – das redes sociais. Se aglomeram e se aproximam aqueles com maior afinidade. A repulsão é uma força que afasta indicando menor afinidade. Investimos nessa abordagem, pois os módulos apresentam visualmente o espaço de interesse para ação dos *bots* em modelos de propaganda em rede.

O último passo para moldar uma perspectiva é explorar a seção "Laboratório de Dados" do Gephi, um quadro semelhante a uma tabela do Excel. Nessa tabela o pesquisador deve exportar os nomes indicados em cada *cluster*, gerando um novo arquivo separado por vírgulas referente àquela perspectiva. O procedimento continua a partir desse novo arquivo, que chamaremos de *clusterusernames.csv*, junto com o arquivo inicial carregado com todos os *clusters* denominado *tweets.csv*. Compreendendo que as partes são maiores que o todo, isolamos, com esse último processamento, os *tweets* e metadados que cada *cluster* carrega consigo. Se cada *cluster* realizar um novo processamento do algoritmo FORD, verá que será possível gerar um novo arcabouço de informações relacionais referentes às expressões que modelam cada perspectiva. Podem ser informações como a rede de *hashtags*, palavras mais frequentes, usuários mais ativos e *hashtags* mais frequentes. As possibilidades são variadas em termos de métricas.

Depois de identificarmos o cenário de expressão nas perspectivas, aplicamos a estatística de grau ponderado para medir volume de *tweets* e retweets. Para analisar a influência dos perfis na rede e detectar anomalias no comportamento dos usuários, processamos essa estatística para visualizar aqueles nós com maior ou menor grau de *tweets*. O grau ponderado médio usa a soma média dos pesos das arestas conectadas a um nó. Essa métrica pode indicar características de táticas de spam. O grau ponderado fornece as métricas de saída e entrada de um nó, medindo o volume de mensagens que

divulgar o debate televisivo nas redes sociais. Percebemos a proximidade da rede aecista com as tags da TV, enquanto isso, a comunidade de tags dilmistas conserva uma certa distância. Por termos feito também do dia posterior à votação do segundo turno (28-10-14), a hashtag #RIPBrasil acabou ganhando destaque diante da indignação dos eleitores de Aécio Neves (PSDB) e da rede antipetista.

Tabela 2 – Com os tweets mais compartilhados durante os 8 dias (21-28 de outubro) de coleta das eleições presidenciais de 2014

Tweet	Retweets
RT @g1: Dilma chega	48016
RT @humdaora: no caso eu sou a dilma http://t.co/vDfdJiNqBA	9804
RT @titwo: fanfic: um dia antes da eleição, dilma e aécio acordam em corpos trocados. agora cada um tem 24h para convencer o brasil a não ...	7690
RT @Sirjoker_: MANO DO CÉU A DILMA TA MUITO MAGRA http://t.co/Q5n1KVYCyS	6598
RT @adamschmidel: Eu, você, dois filhos, férias em paris, mansão em miami, internet de 100 megas e 50 milhões na conta ♥	6327
RT @rodpocket: “o aécio sabe falar melhor que a dilma” se eu quisesse alguém que sabe falar eu votava no bial pois o video do filtro solar ...	5730
RT @malikdefenses: Dilma: 13 Marina: 40 Aécio: 45 CPM: 22 Ben: 10 Onze: 20 Nx: 0 Marcelo D: 2 U: 2 P: 9 Blink: 182 Brasil: 1 Alemanha: 7	5487
RT @barbiesemken: Tudo isso pra chegar hoje e votar na dilma... http://t.co/0UqOcsN54a	4320
RT @larrible: aecio disse que vai cuidar pessoalmente da segurança ou seja você estara dormindo e ele vai passar na rua de noite fazendo ro...	3896
RT @10Ronaldinho: O Brasil já mostrou que tem potencial para crescer muito mais!! Domingo vamos para as urnas mudar nosso país!! #Aecio45 #...	3836

Fonte: Twitter

Essa análise a partir dos dados de oito dias de coleta sobre as eleições para a presidência do Brasil em 2014 tem como objetivo entrar no ambiente do jogo político em formação na disputa eleitoral. O método quali-quantitativo utiliza dos dados e da observação (tanto visual, quanto de leitura), para construir o plano de pesquisa transversal à computação e à comunicação. O humor e a ironia acabam gerando um alto grau de compartilhamento. Um comportamento robótico não tem a capacidade de criar um tweet emocionalmente comprometido com o tema, mas essa mensagem pode ser repetidamente compartilhada.

Os novos dados nos entregam uma tabela com os usuários que mais publicaram durante o período. Em nosso percurso de investigação, a revisão de literatura e as pistas para a metodologia, o número de publicações em curto espaço de tempo indica graus de automação no comportamento do perfil. No quadro a seguir, os dez usuários mais ativos durante a última semana das eleições presidenciais apresentaram características de automação como alto número de publicações, retweetam ou são retweetados com intensidade alta, alto número de menções com baixo grau de respostas. Alguns deles já estão em hiato (quando um perfil fica um tempo sem publicar); outro publicou bastante, mas pouco interagiu; um usuário não utiliza foto de perfil; outros foram apenas retweetados; a maioria foi mencionada com alta frequência. Nota-se, nesses perfis, uma tentativa de criar uma personalidade para tal usuário. No processo de automação desses usuários – militância, marketing pago, bots, fãs eleitores, perfis “alugados” – pode-se apontar ao observar as páginas dos perfis um engajamento visual e léxico de expressão no jogo político eleitoral.

Tabela 3 – Com os dez usuários mais ativos durante o período de 21 a 27 de outubro de 2014. A coluna tweets contém as mensagens publicadas por esses usuários em ordem decrescente

Perfil	Tweets	Retweetou	Retweetado	Respondeu	Respondido	Mencionou	Mencionado
raynnerem	3308	6	3	0	0	1	3
dairosoares	2471	165	2362	6	14	80	1100
letrin45	2424	162	2213	17	142	48	616
fmoreira66	2412	0	2401	5	10	96	860
artur_jaru	2364	1	2354	3	7	7	755
tottustu	2240	3844	502	183	1451	1338	2482
rfgv63	2195	0	2191	2	4	8	470
keunecke	2058	652	1958	20	3	167	722
allansb	1941	310	866	44	63	55	307
aokadaokada	1835	73	1813	1	1	18	490

Fonte: Twitter

Com os grafos visuais e as métricas sobre os comportamentos dos perfis podemos mapear as pulsações na rede. Na prática, as métricas estatísticas junto com uma análise qualitativa dos tweets equalizam a metodologia quantitativa-qualitativa da pesquisa com dados sociais. O contexto topológico e temporal dos sinais desses efeitos coordenados no grafo indica o modo de ação estratégica das *botnets*. Os usuários mais ativos servem de perfis âncoras para nossa análise, mas não só. As redes de *bots* – como visto nos trabalhos das ciências da computação – funcionam *in loco* para potencializarem determinadas *hashtags*. Uma metáfora interessante seria imaginar que as *botnets* se assemelham ao uso do doping por atletas no esporte. O doping caracteriza-se pelo uso de substâncias que podem alterar a resposta do corpo frente a um estímulo. Em geral, os casos de doping revelam que os atletas em uma determinada competição pretendem com o uso de substâncias potencializar seu rendimento, força, agilidade ou até mesmo perda de peso. Essa relação entre as estatísticas de retweets, usuários mais ativos, modularidade dependem da contextualização do evento ou objetivo do estímulo. Nas eleições presidenciais de 2014, as campanhas de ambos os candidatos utilizam estrategicamente as redes de *botnets* para contaminar as *hashtags* #Aecio45PeloBrasil e #SomosTodosDilma para potencializar as cascatas de informação entre seus seguidores durante os debates e nos últimos dias da campanha eleitoral. Entretanto, isso não significa – como nós sabemos sobre atletas que “caem” no exame antidoping – que o rendimento ou a circulação da interação esteja ocupada totalmente por contas automatizadas. Ou ainda, que a trajetória dessas conversas seja feita do início ao fim por redes de *bots*. Os *bots* como agentes de expressão operacionalizam uma estratégia de contaminação automatizada em cascata com o poder de criar redes de interação com humanos. Quando utilizado em plataformas de interação on-line, o objetivo é potencializar e modelizar o comportamento das interações entre as conversas e pessoas durante um espaço de tempo.

Considerações finais: Bots como agentes de expressão

Apesar de reconhecerem as redes sociais como uma camada extensiva do espaço público democrático de conversação, as pesquisas vindas do campo da computação priorizam o desenvolvimento de métodos que demonstram a vulnerabilidade dos sites de redes sociais, isto é, o quanto estão suscetíveis à circulação e aos ataques de bots. Pesquisas no campo da ética em inteligência artificial e da informação (FLORIDI, 2019; DURANTE, 2019) indicam a necessidade da continuidade

da exploração de abordagens que considerem agentes informacionais e suas responsabilidades, propondo modelos de algoritmos para detectar atividades automatizadas e seus efeitos graduais em um sistema complexo nos mais diversos ambientes, como a biomedicina, serviços públicos de gestão e administração, de ordenamento jurídico e, é claro, a comunicação. Taddeo e Floridi (2018) propõem um programa ético translacional para a inteligência artificial, voltado à formulação de metodologias de prospecção que possam indicar riscos e oportunidades, com o objetivo de evitar consequências indesejadas. Na prática eles propõem análises de avaliação de impacto em todos os passos de implementação de práticas e tecnologias de inteligência artificial, considerando aspectos como privacidade, transparência ou responsabilidade.

As eleições para a presidência da república em 2014 demonstraram como a circulação de informações na internet tem relevância para a formação da opinião pública, bem como para as decisões dos eleitores. Entretanto, quando se fala de jogos de poder, o algoritmo tem acirrado o debate ético sobre até que ponto uma ferramenta pode modular estruturalmente conversas ao ponto de interferir no jogo democrático. O uso de estratégias robóticas – rotinas computacionais ou a imitação do comportamento automatizado – pode implantar um tipo de verdade (BEER, 2016). Bruno e Vaz (2002) já apontaram ao analisar a tecnologia de recomendação da Amazon.com que humanos e tecnologias cognitivas compartilham da distribuição da inteligência. Esse *agente.com* não é uma prótese da humanidade e suas tarefas não se limitam ao prolongamento de funções cognitivas previamente registradas, mas sim um objeto técnico operando com “uma diferença, um desvio, uma transformação na atividade cognitiva” (BRUNO; VAZ, 2002, p. 36). Os *bots* como agentes de expressão se organizam em *botnets* coordenando tarefas, distribuindo-se na rede de acordo com metas específicas e globais. De acordo com o cenário e a situação, os *bots* atuam gerenciando modularidades e ativando padrões de ativação no fluxo de informação nas redes sociais. Essa investigação futuramente pretende explorar noções clássicas de inteligência artificial sobre conhecimento, aprendizado, planejamento e adaptação para analisar os efeitos dessas tecnologias nos estudos comunicacionais.

Referências

- ALLEN, Colin; SMIT, Iva; WALLACH, Wendell. Artificial morality: Top-down, bottom-up, and hybrid approaches. **Ethics and Information Technology**, v. 7, n. 3, p. 149-155, 2005. Disponível em: https://www.researchgate.net/publication/225850648_Artificial_Morality_Top-down_Bottom-up_and_Hybrid_Approaches. Acesso em: 12 fev. 2015.
- ARNAUDO, Dan. Computational Propaganda in Brazil: Social Bots During Elections. **Computational Propaganda Project Working Paper Series**, v. 8, 2017. Disponível em: <https://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Brazil-1.pdf>. Acesso em: 15 mar. 2018.
- BASTIAN, Mathieu; HEYMANN Sebastian; JACOMY, Mathieu. Gephi: an open source software for exploring and manipulating networks. **3rd International AAI Conference on Weblogs and Social Media**, San Jose Marriott Hotel, 17 a 20 de maio de 2009. Disponível em: <https://gephi.org/publications/gephi-bastian-feb09.pdf>. Acesso em: 16 abr. 2017.
- BENEVENUTO, Fabricio et al. Detecting Spammers on Twitter. **7th Annual Collaboration, Electronic messaging, AntiAbuse and Spam Conference**, Microsoft's Main Campus in Redmond, 13 a 14 de julho de 2010. Disponível em: <https://homepages.dcc.ufmg.br/~fabricio/download/ceas10.pdf>. Acesso em: 12 out. 2017.
- BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. **Network propaganda: manipulation, disinformation, and radicalization in American Politics**. Oxford, UK: Oxford University Press, 2018.
- BEER, David. The social power of algorithms. **Information, Communication & Society**, v. 20, n. 1, 2017, p. 1-13. Disponível em: <https://doi.10.1080/1369118X.2016.1216147>. Acesso em: 20 ago. 2018.

BLONDEL, Vincent et al. **Fast unfolding of communities in large networks**. In: *Journal of Statistical Mechanics: Theory and Experiment*: v. 10, 2008. Disponível em: <https://arxiv.org/abs/0803.0476>. Acesso em: 3 jun. 2016.

BRUNO, Fernanda; VAZ, Paulo. Agentes.com: cognição, delegação, distribuição. **Contracampo**, v. 7, n. 1, p. 23-38, 2008. Disponível em: <http://periodicos.uff.br/contracampo/article/view/17333>. Acesso em: 16 set. 2018.

CANUTO, Sergio; GONÇALVES, Marcos Andre; BENEVENUTO, Fabricio. Exploiting new sentiment- based meta-level features for effective sentiment analysis. In: **Proceedings of the Ninth ACM International Conference on Web Search and Data Mining**, p. 53-62, 2016. Disponível em: <https://dl.acm.org/citation.cfm?id=2835821>. Acesso em: 23 dez. 2018.

CHAKRABORTY, Abhijnan et al. Who Makes Trends? Understanding Demographic Biases in Crowdsourced Recommendations. **11th International AAAI Conference on Web and Social Media**, Hyatt Regency Montreal, 16 a 18 de maio de 2017. Disponível em: <https://arxiv.org/abs/1704.00139>. Acessado em: 18 jan. 2018.

CÔRTEZ, Thaisa Guimaraes; ZIGONI, Luísa Perdigão; CANCIAN, Allan; MALINI, Fabio Luiz. O #VemPraRua em dois ciclos: análise e comparação das manifestações no Brasil em 2013 e 2015. **XXXIX Congresso Brasileiro de Ciências da Comunicação**, Universidade de São Paulo, 5 a 9 de setembro de 2016. Disponível em: <http://portalintercom.org.br/anais/nacional2016/resumos/R11-1938-1.pdf>. Acesso em: 27 fev. 2017.

CHENG, Justin et al. Predicting reciprocity in social networks. **Anais do Third International Conference on Social Computing**, p. 49-56, 2011. Disponível em: <https://ieeexplore.ieee.org/document/6113094>. Acesso em: 8 set. 2017.

CURTIS, Pavel. Mudding: social phenomena in text-based virtual realities. In: LUDLOW, Peter (Ed.). **High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace**. Cambridge, MA: MIT Press, 1996. Disponível em: <https://pdfs.semanticscholar.org/79b8/93a15ea0d6bb4cb59fe197423b00eef92277.pdf>. Acesso em: 19 mar. 2018.

DURANTE, Massimo. **Ethics, Law and the Politics of Information – A Guide to the Philosophy of Luciano Floridi**. [The International Library of Ethics, Law and Technology Series]. New York: Springer Netherlands, 2017.

ESTRADA, Matias; MENDOZA, Marcelo. Affinity prediction in online social networks. **6th Chilean Conference on Pattern Recognition**, Universidad de Talca, 10 a 14 de novembro de 2014. Disponível em: <https://arxiv.org/abs/1408.2871>. Acesso em: 22 out. 2018.

FERRARA, Emilio et al. The rise of social bots. **COMMUNICATIONS OF THE ACM** |59, p. 96-104, 2016. Disponível em: <https://m-cacm.acm.org/magazines/2016/7/204021-the-rise-of-social-bots/fulltext?mobile=true>. Acesso em: 19 jul. 2018.

FERRARA, Emilio. Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election. **First Monday**, v. 22, n. 8, p. 1-33, 2017. Disponível em: <https://arxiv.org/pdf/1707.00086.pdf>. Acesso em: 15 jun. 2018.

FLORIDI, Luciano. **The Logic of Information – A Theory of Philosophy as Conceptual Design**. Oxford, UK: Oxford University Press, 2019.

FORTUNATO, Santo. Community detection in graphs. **Physics Reports** 486, p. 75-174, 2010. Disponível em: <https://arxiv.org/abs/0906.0612>. Acesso em: 10 jul. 2017.

FREITAS, Carlos et al. Reverse engineering socialbot infiltration strategies in twitter. **Anais do IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining**, p. 25-32, 2015. Disponível em: https://socialnetworks.mpi-sws.org/papers/TwitterBots_ASONAM15.pdf. Acesso em: 11 ago. 2017.

FREITAS, Carlos. Socialbots: Implicações na segurança e na credibilidade de serviços baseados no twitter. **Anais do 32º Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**, p. 603-616, 2014. Disponível em: <http://sbrc2014.ufsc.br/anais/files/trilha/ST14-2.pdf> Acesso em: 16 maio 2016.

GUATTARI, Felix. A Propósito de las Máquinas. In: NADAUD, Stéphane. (Org.) **¿Que és la Ecosofia?**. Buenos Aires: Cactus, 2015. p. 87-101.

GILLESPIE, Tarleton. #Trendingistrending: When Algorithms Become Culture. In: SEYFERT, Robert; ROBERGE, Jonathan (Eds.) **Algorithmic Cultures: Essays on Meaning, Performance, and New Technologies**. London: Routledge, 2017. Disponível em: <https://www.microsoft.com/en-us/research/publication/trendingistrending-when-algorithms-become-culture-3/>. Acesso em: 13 fev. 2018.

HOWARD, Philip N.; KOLLANY, Bence. Bots, #strongerin, and #brexit: Computational propaganda during the UK-EU referendum. **The Computational Propaganda Program**, 21 jun. 2016. Disponível em: <https://comprop.oii.ox.ac.uk/research/working-papers/bots-strongerin-and-brexit-computational-propaganda-during-the-uk-eu-referendum/>. Acesso em: 11 mar. 2017.

HOWARD, Philip; WOOLEY, Samuel; CALO, Ryan. Algorithms, bots, and political communication in the US. 2016 election: The challenge of automated political communication for election law and administration. **Journal of Information Technology & Politics**, v. 15, n. 2, p. 81-93, 2018. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/19331681.2018.1448735>. Acesso em: 19 jul. 2018.

JAMIESON, Kathleen Hall. **Cyberwar**: how russian hackers and trolls helped elect a president. HighBridge Audio, 2018.

LEONARD, Andrew. **BOTS: The origin of New Species**. England: Penquin Books, 1997.

MAES, Patti (Org.). **Designing Autonomous Agents**. Massachusetts: MIT Press, 1990.

MAGALHÃES, João Carlos. Do Algorithms Shape Character? Considering Algorithmic Ethical Subjectivation. **Social Media + Society**, apr. 2018. Disponível em: doi:10.1177/2056305118768301. Acesso em: 12 jan. 2019.

MALINI, Fabio. Um método perspectivista de análise de redes sociais. **Relatório do Grupo de Trabalho Comunicação e Cibercultura**. Associação Nacional dos Programas de Pós-Graduação em Comunicação de Goiânia, 2016. Disponível em: http://www.labic.net/wp-content/uploads/2016/06/compos_Malini_2016.pdf. Acesso em: 12 fev. 2017.

MARWICK, Alice; LEWIS, Rebecca. Media Manipulation and Disinformation Online. **Data & Society report**. Data & Society Research Institute, 2018. Disponível em: <https://datasociety.net/output/media-manipulation-and-disinfo-online/>. Acesso em: 13 jan. 2019.

MESSIAS, Johnnatan et al. White, Man, and Highly Followed: Gender and Race Inequalities in Twitter. **IEEE/WIC/ACM International Conference on Web Intelligence**, Leipzig, 23 a 27 de agosto de 2017. Disponível em: https://homepages.dcc.ufmg.br/~fabricio/download/WI_White_Man_Highly.pdf. Acesso em: 18 mar. 2018.

NADLER, Anthony; CRAIN, Mathew; DONOVAM, Joan. Weaponizing the Digital Influence Machine The Political Perils of Online Ad Tech. **Data & Society report**. Data & Society Research Institute, 2018. Disponível em: https://datasociety.net/wp-content/uploads/2018/10/DS_Digital_Influence_Machine.pdf. Acesso em: 16 jan. 2019.

PATODKAR, Vaibhavi N. et al. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 5, n. 12, dec. 2016. Disponível em: <https://www.aclweb.org/anthology/L10-1263/>. Acesso em: 17 mar. 2018.

PHILLIPS, Whitney; MILNER, Ryan. **The Ambivalent Internet: Mischief, Oddity, and Antagonism Online**. Cambridge, UK: Polity Press, 2017.

PHILLIPS, Whitney. The Oxygen of Amplification Better Practices for Reporting on Extremists, Antagonists, and Manipulators. **Data & Society Report**. Data & Society Research Institute, 2018. Disponível em: <https://datasociety.net/output/oxygen-of-amplification/>. Acesso em: 19 jan. 2019.

REGATTIERI, Lorena. **Guerra de mundos**: a estratégia robótica no twitter e as eleições presidenciais de 2014. Dissertação (Mestrado em Comunicação Social) – Universidade Federal do Espírito Santo, 2016. Disponível em: http://bdtd.ibict.br/vufind/Record/UFES_07aebd33238bd65dabdc2e035243906e. Acesso em: 15 jan. 2018.

REGATTIERI, Lorena et al. Marco Civil: Visualizing the Civil Rights Framework for the Internet in Brazil. **HT CEUR Workshop Proceedings**, Chile, 2014. Disponível em: https://www.researchgate.net/publication/289664415_MarcoCivil_Visualizing_the_Civil_rights_Framework_for_the_Internet_in_Brazil. Acesso em: 16 jan. 2018.

SANGER, David. **The Perfect Weapon**: war, sabotage, and fear in the cyber age. Danvers: Crown Publishing Group, 2018.

SAVVOPOULOS, Alkiviadis.; PANTELIS, Vikato; BENEVENUTO, Fabricio. Socialbots' First Words: Can Automatic Chatting Improve Influence in Twitter? **IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)**, Castell Jalpí, 28 a 31 de agosto de 2018. Disponível em: <https://homepages.dcc.ufmg.br/~fabricio/download/asonam2018-socialbots.pdf>. Acesso em: 15 jan. 2019.

SHAO, Chengcheng et al. The spread of low-credibility content by social bots. **Nature Communications**, v.9, 2018. Disponível em: <https://www.nature.com/articles/s41467-018-06930-7>. Acesso em: 18 dez. 2018.

SILVA, Leandro et al. Analyzing the Targets of Hate in Online Social Media. **10th International AAAI Conference on Weblogs and Social**, Leibniz Institute for the Social Sciences, 16 a 20 de maio de 2016. Disponível em: <https://homepages.dcc.ufmg.br/~fabricio/download/icwsm2016-hate.pdf>. Acesso em: 25 jun. 2017.

SHOREY, Samantha; HOWARD, Philip. Automation, Big Data, and Politics: A Research Review. **International Journal of Communication**, v. 10, p. 5032–5055, 2016. Disponível em: <https://ijoc.org/index.php/ijoc/article/view/6233>. Acesso em: 18 jul. 2017.

TADDEO, Mariarosaria; FLORIDI, Luciano. How AI can be a force for good: An ethical framework will help to harness the potential of AI while keeping humans in control. **Science**, v. 361, n. 6404, p. 751-752, aug. 2018. Disponível em: 10.1126/science.aat5991. Acesso em: 17 jul. 2019.

TURKLE, Shirley. Constructions and reconstructions of self in virtual reality: Playing in the MUDs. **Mind, Culture, and Activity**, v. 1, n. 3, p. 158-167, 1994. Disponível em: http://web.mit.edu/sturkle/www/pdfsforstwebpage/ST_Construc%20and%20reconstruc%20of%20self.pdf Acesso em: 16 jun. 2017.

TURING, Alan. Computing Machinery and Intelligence. In: BODEN, Margaret A. (Ed.) **The philosophy of Artificial Intelligence**. Oxford, UK: Oxford University Press, 1990. Disponível em: <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>. Acesso em: 26 out. 2018.

VAROL, Onur et al. Online human-bot interactions: detection, estimation, and characterization. **11th International AAAI Conference on Web and Social Media**, Hyatt Regency Montreal, 16 a 18 de maio de 2017. Disponível em: <https://arxiv.org/abs/1703.03107>. Acesso em: 12 fev. 2018.

VOSOUGUI, Soroush; ROY, Deb; ARAL, Sinan. The spread of true and false news online. **Science**, v. 359, p. 1146–1151, 2018. Disponível em: <http://ide.mit.edu/sites/default/files/publications/2017%20IDE%20Research%20Brief%20False%20News.pdf> Acesso em: 20 jan. 2019

WAGNER, Claudia et al. When social bots attack: Modeling susceptibility of users in online social networks. **2nd Workshop Making Sense of Microposts (#MSM2012)**, 2012. Disponível em: http://ceur-ws.org/Vol-838/paper_11.pdf. Acesso em: 18 fev. 2017.

WOOLEY, Samuel; KOLLANYI, Bencer; HOWARD, Philip. Bots and Automation over Twitter during the U.S. Election. **COMPROP DATA MEMO**, v. 4, 2016. Disponível em: <https://comprop.oi.ox.ac.uk/research/working-papers/bots-and-automation-over-twitter-during-the-u-s-election/>. Acesso em: 12 mar. 2017.