



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Amanda Maria de Almeida Nunes

MÁQUINAS SOCIAIS E A DESINFORMAÇÃO EM REDE: O papel das entidades de *software* na formação de opinião na Internet

Recife

2020

Amanda Maria de Almeida Nunes

MÁQUINAS SOCIAIS E A DESINFORMAÇÃO EM REDE: O papel das entidades de *software* na formação de opinião na Internet

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Área de Concentração: Informação, Memória e Tecnologia.

Orientador: Prof. Dr. Célio Andrade de Santana Júnior.

Recife

2020

Catálogo na fonte
Bibliotecária Andréa Carla Melo Marinho, CRB-4/1667

N972m Nunes, Amanda Maria de Almeida
Máquinas sociais e a desinformação em rede: o papel das entidades de *software* na formação de opinião na internet / Amanda Maria de Almeida Nunes. – Recife, 2020.
175f.: il.

Orientador: Célio Andrade de Santana Júnior.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Programa de Pós-Graduação em Ciência da Informação, 2020.

Inclui referências.

1. Máquinas sociais. 2. Desinformação. 3. Entidades não humanas. 4. Notícias falsas.
I. Santana Júnior, Célio Andrade de (Orientador). II. Título.

020 CDD (22. ed.)

UFPE (CAC 2020-147)

Amanda Maria de Almeida Nunes

Máquinas sociais e desinformação em rede: o papel das entidades de software na formação de opinião na internet

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de mestre em Ciência da Informação.

Aprovada em: 13/02/2020

BANCA EXAMINADORA

Prof. Dr. Célio Andrade de Santana Júnior (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Fábio Mascarenhas e Silva (Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Fábio Castro Gouveia (Examinador Externo)
Fundação Oswaldo Cruz

*Ao meu pai, Valdir Nunes de Freitas e a
minha mãe Maria Solange de Almeida Nunes,
que depositaram em mim sua fé e esperança,
meu muitíssimo obrigada, vocês são os
alicerces da minha trajetória e sucesso.*

AGRADECIMENTOS

Hoje encerro mais uma etapa da minha formação, me despeço da Universidade Federal de Pernambuco (UFPE) não apenas com o título de Mestre, mas com a certeza de que vivi e cresci extraordinariamente durante estes dois anos.

Por isso, agradeço imensamente a Deus que na sua infinita bondade me permitiu chegar até aqui com sabedoria e respeito a todos que encontrei durante a minha trajetória. Também quero agradecer a mim mesma por não ter desistido de lutar, por mais do que me permitiram ter, por não ter fraquejado ou me acomodado ao longo da estrada.

A minha família obrigada de coração e a todos vocês que fazem parte do que eu sou, para onde eu vou e de onde eu venho. Agradeço por depositarem em mim suas apostas de confiança, respeito e admiração, espero ter honrado vocês ao longo desses anos e por todos os próximos que virão.

Aos amigos que fiz durante o processo, saibam que vocês foram essenciais para aliviar o peso que uma pós-graduação traz, foi com vocês que descobri ser possível fazer pesquisa com sorriso no rosto e satisfação, levo todos comigo e exemplo para a vida, espero ter também contribuído de alguma forma para o crescimento de vocês.

Registro aqui minha infinita gratidão à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por todo apoio e infraestrutura oferecida por meio da bolsa de Mestrado; aproveito a oportunidade para fazer um apelo em nome da produção científica brasileira, para que seja possível outros jovens, assim como me foi permitido, gozarem do benefício e incentivo oferecido por essa instituição.

Ao meu orientador Célio Andrade de Santana Júnior meus mais sinceros agradecimentos pela dedicação, esforço, companheirismo e ética, a sua condução foi essencial para o sucesso alcançado nesta pesquisa e as lições que aprendi ultrapassam os resultados que aqui estão expostos.

Ao Departamento de Ciência da Informação e a Universidade Federal de Pernambuco, seus funcionários e professores, agradeço a acolhida, e saibam que levo comigo as melhores recordações e o maior tesouro que se pode ter na vida, o CONHECIMENTO.

A falsidade voa e a verdade vem mancando depois disso, de modo que quando os homens se tornam indecisos, é tarde demais; a brincadeira acabou e o conto teve seu efeito (SWIFT, 1710).

RESUMO

Apresenta o contexto atual, onde máquinas sociais participam do processo de desinformação na *Web*, influenciando o debate através do compartilhamento massivo de informações inverídicas, enviesadas ou distorcidas. O objetivo deste trabalho é analisar a participação das máquinas sociais (*software*) no processo de desinformação que ocorre nesse ambiente, em que já são suficientemente autônomas para elaborar estratégias próprias, com a finalidade de atingir suas metas. O método da pesquisa é um estudo de caso, para tanto destaca-se o contexto das eleições brasileiras de 2018, onde observa-se a participação das máquinas sociais na propagação da desinformação que trouxe repercussões à sociedade. Desta forma, verifica-se que as máquinas sociais vêm sendo utilizadas para disseminar informações neste contexto, observando-se que através do acesso a grande quantidade de informação (*big data*), tais máquinas estão sendo mais seletivas e efetivas no processo de desinformação. Ressalta-se que a influência promovida pelas máquinas pode acontecer de forma localizada, sugerindo o comportamento de uma pessoa ou até mesmo, em eventos maiores como as eleições e os debates sociais, que atinjam grande repercussão nas redes sociais virtuais. Conclui-se que, com a ampliação da adoção da Internet no cotidiano das pessoas, cada vez mais as máquinas sociais farão a seleção/curadoria do conteúdo a ser apresentado aos usuários e, conseqüentemente, o aumento da exposição dos mesmos a informações que expõem as pessoas a situações que desencadeiam processos de desinformação, deve-se considerar também a especialização dessas ferramentas e a sua sofisticação no processo de desinformar, induzindo comportamentos através das mensagens que dissemina, iniciando um poderoso processo de manipulação em contextos políticos, econômicos ou sociais.

Palavras-Chave: Máquinas Sociais. Desinformação. Entidades não humanas. Notícias Falsas.

ABSTRACT

It presents the current context, where social machines participate in the process of disinformation on the web, influencing the debate through the massive sharing of untrue, biased or distorted information. The objective of this work is to analyze the participation of social machines (*software*) in the disinformation process that occurs in this environment in which they are already sufficiently autonomous to develop their own strategies in order to achieve their goals. The research method is a case study, therefore, the context of the Brazilian elections of 2018 stands out, where the participation of social machines in the propagation of the disinformation that has had repercussions on society is observed. Thus, it appears that social machines have been used to disseminate information in this context, observing that through access to a large amount of information (big data), such machines are being more selective and effective in the disinformation process. It is noteworthy that, the influence promoted by the machines can happen in a localized way, suggesting a person's behavior or even in larger events such as elections and social debates that reach great repercussions in virtual social networks. It is concluded that, with the expansion of the adoption of the Internet in people's daily lives, more and more social machines will make the selection / curation of the content to be presented to users and, consequently, their exposure to information that exposes people to situations that trigger disinformation processes, the specialization of these tools and their sophistication in the disinformation process must also be considered, inducing behaviors through the messages it disseminates, initiating a powerful manipulation process in political, economic or social contexts.

Keywords: Social Machines. Misinformation. Nonhuman entities. False News.

LISTA DE FIGURAS

Figura 1 - Máquinas Sociais e seus Componentes.....	99
Figura 2 - Comparação entre as entidades.....	106
Figura 3 - Relação entre as entidades.....	114
Figura 4 - Etapas da Pesquisa.....	124
Figura 5 - Origem de contas suspeitas apoiando Aécio Neves	130
Figura 6 - O Facebook menciona em 28 de setembro de 2014.....	131
Figura 7 - Monitoramento do Twitter durante o Debate da Globo (1º Turno)...	132
Figura 8 - Monitoramento do Twitter durante o Debate da Globo (2º Turno)...	134
Figura 9 - Atividade no Twitter relacionada ao Impeachment	137
Figura 10 - Concentração de Contas Suspeitas (Impeachment)	140
Figura 11 - Interconexões de usuários no grupo de eleições do WhatsApp	142

LISTA DE QUADROS

Quadro 1 - Tipos de Desinformação por Skyrms	44
Quadro 2 - Resumo das Facetas da Desinformação	72

SUMÁRIO

1	INTRODUÇÃO	14
2	DESINFORMAÇÃO	22
2.1	Desinformação como Coisa, Processo e Propósito	31
2.1.1	Desinformação como Coisa	32
2.1.2	Desinformação como Processo.....	37
2.1.3	Desinformação como Propósito	43
2.2	Manifestações da Desinformação.....	54
2.2.1	Pós-Verdade	54
2.2.2	Notícias falsas	56
2.2.3	Data Flood (Enxurrada de Dados).....	59
2.2.4	Falsa Propaganda	60
2.2.5	<i>Internet Bias</i> (Viés da Internet).....	61
2.2.6	Social Spam	62
2.2.7	Deepfake	63
2.3	Movimentos Criados a partir da Desinformação	65
2.3.1	Vigilantismo	65
2.3.2	Checagem de Fatos	66
2.3.3	Tecnocensura.....	68
2.3.4	Filtro Bolha	70
2.4	Síntese da Seção.....	71
3	MÁQUINAS SOCIAIS E ENTIDADES NÃO BIOLÓGICAS	74
3.1	Robôs	74
3.2	Máquinas.....	77
3.3	Autômatos.....	80

3.4	Androides.....	82
3.5	Ciborgues.....	83
3.6	Agentes de <i>Software</i>.....	85
3.7	Bots	88
3.8	Máquinas Sociais	92
3.9	Comparação entre as Entidades.....	102
3.10	Síntese da Sessão.....	117
4	PROCEDIMENTOS METODOLÓGICOS	119
4.1	Execução da Pesquisa.....	124
4.1.1	Seleção do Estudo de Caso.....	125
4.1.2	Coleta de Dados.....	126
4.1.3	Análise de Dados	127
5	DESINFORMAÇÃO NO CENÁRIO POLÍTICO DO BRASIL (2014-2018)	129
5.1	Eleições 2014.....	129
5.1.1	Exército de Máquinas Sociais	130
5.1.2	Participação das Máquinas Sociais.....	131
5.1.3	Desinformação (<i>hoax</i>).....	135
5.1.4	Resumo das eleições de 2014	136
5.2	Impeachment 2016	136
5.2.1	Resumo do Impeachment	141
5.3	Eleições 2018.....	141
5.3.1	Resumo das Eleições 2018.....	144
6	CONCLUSÃO	145
6.1	Descobertas.....	145
6.2	Discussão dos Resultados.....	150

6.3	Ameaças a Validade	154
6.4	Impactos Sociais	157
6.5	Trabalhos Futuros	159
	REFEFÊNCIAS	162

1 INTRODUÇÃO

A popularização do uso da Internet nas diversas camadas da sociedade civil trouxe mudanças significativas no comportamento das pessoas face à larga adoção dessa tecnologia. Esse cenário, proporcionou o surgimento de contextos distintos no que diz respeito ao comportamento do usuário no processo de produção e disseminação da informação, bem como na recuperação/recepção de conteúdo. Destaca-se que a partir da metade da década de 1990 e início dos anos 2000 verificou-se a explosão do uso doméstico da Internet, chegando ao número de um bilhão de lares conectados, de acordo com relatório da “*ICT Facts & Figures*” (ITU, 2016, p. 4).

No Brasil, entre 2005 e 2015, é possível observar uma expansão do acesso à Internet, que passou dos 7,2 milhões para 39,3 milhões, o que representa um aumento médio de 446% no período, conforme destaca a Pesquisa Nacional por Amostragem de Domicílio, realizada pelo Instituto Brasileiro de Pesquisa e Estatística (IBGE, 2016, p. 81).

Nessa perspectiva, considera-se o aumento do uso de celulares, *tablets* e outros equipamentos, além do computador tradicional, como meio de acesso à rede. Isso sugere a alteração de costumes não apenas no que diz respeito ao uso e manipulação desses suportes tecnológicos, mas do próprio envolvimento do usuário para com a produção e uso da informação nesse ambiente.

Dentro dessa realidade, surgem as redes sociais que se mostram como ‘espaços’ de compartilhamento de dados sociais, e alguns destes, tais como o *Facebook*, o *Twitter*, e o *Whatsapp*, atingiram rapidamente a marca de 1 bilhão de usuários. Portanto, verifica-se que os espaços de compartilhamento (do mundo ‘real’) migraram para o ambiente digital, onde os seres humanos veem potencializadas a sua capacidade de comunicação, simplesmente, por estarem e atuarem em rede.

A expansão dessas relações virtuais se deu, principalmente, pela crescente utilização de *smartphones* combinados com a Internet, que, além de potencializar, transforma o modo como as pessoas se comunicam, modifica os regimes de trabalho e até mesmo cria novos hábitos de consumo. Isso resulta na intensa disponibilização de dados pessoais na Internet, denominados dados sociais, justamente por tratar de

conteúdo que caracteriza o perfil de um usuário por suas singularidades, como por exemplo, gostos musicais, culturais, políticos, gastronômicos e afins. Sobre esse aspecto, Doneda, afirma que:

A dinâmica de funcionamento de uma rede social *online* prevê que seus usuários se apresentem perante os demais por meio de suas informações pessoais - sejam estes dados identificativos, gostos, opiniões, mensagens, fotografias, vídeos - enfim, basicamente todos os aspectos de uma vida passíveis de encontrarem expressão em formato digital (DONEDA, 2012, p. 6).

Lampe, Ellison e Steinfield (2006) relatam que se compreende por *software* social aquele que possibilita o encontro, a conexão e a colaboração de pessoas na rede, através de um dispositivo ligado à Internet. A principal característica dessa ferramenta reside no fato de que os usuários utilizam suas plataformas para inserir ideias, crenças, questionamentos, preferências, informações pessoais de uma maneira geral, razão pela qual, boa parte do conteúdo ali gerado são dados sociais. Essa prática não surgiu exclusivamente com as novas redes sociais, que tem um potencial colaborativo maior que as antecessoras, pelo contrário, é algo que existe há anos na forma de fóruns, *blogs* e listas de discussão, mas que ganhou notoriedade a partir da *Web 2.0*.

Verifica-se com isso, o surgimento de diversos sistemas baseados em *softwares* sociais e colaborativos, que trouxeram significativas mudanças na forma de criação, validação e disseminação de conteúdo no cotidiano das pessoas. Angelo (2016, p. 72) afirma que, entre os fenômenos do intenso uso da Internet, pode-se destacar a ampliação da participação nas redes sociais motivadas pela velocidade da comunicação.

Nesses ambientes, há uma série de atores movidos pelos mais variados interesses, que apresentam uma diversidade de participantes em rede, fazendo surgir novos comportamentos informacionais, os quais passaram a ser amplamente divulgados em veículos especializados em tecnologia, publicados na imprensa e em periódicos acadêmicos especializados.

A diversidade aqui referida vai além daquela observada na variedade dos pontos de vista, posições sociais, políticas, econômicas ou de gênero. O que se quer observar é a pluralidade de agentes, que podem ser humanos ou não humanos, envolvidos na disseminação da informação na rede. Aqui, se chama a atenção para a existência de entidades muitas vezes chamadas de Robôs ou Máquinas, que também povoam a rede.

Esses robôs ou máquinas estão participando ativamente das redes sociais, no que diz respeito à interação com outros usuários, inclusive humanos, sendo responsáveis pelo processamento de dados (sociais); neste sentido, ressalta Barth (2018) que 53,1% do tráfego de informações na Internet de 2017 foi realizado por *bots*. Esses números são parecidos com aqueles apresentados por Fischer (2017) o qual sugeria que 51,8% do tráfego de 2016 foi realizado por robôs.

A conjuntura apontada por Barth (2018) e Fischer (2017) preconiza que a participação dessas entidades na rede leva ao caminho inevitável da interação e conseqüentemente a troca de informação entre os usuários, devendo ser observada a capacidade (altíssima) de processamento que esses robôs possuem, facilitando seus processos de aprendizagem e as ações de compartilhamento (disseminação) da informação na Internet.

Se esses agentes participam de maneira tão ativa nos fluxos de informação em rede, inclusive consumindo dados de usuários (humanos) nas redes, pode-se prever que eles também estão envolvidos em processos de desinformação, que é o objeto de estudo deste trabalho, a partir da perspectiva das máquinas sociais, considerando essas como agentes ativos e atuantes nesse fenômeno. Assim, sugere-se que existem máquinas capazes de processar dados rapidamente e interagir com os seres humanos.

Por esse aspecto, discute-se o termo 'máquinas sociais' que de maneira geral podem ser entendidas como uma espécie evoluída dos agentes de *software* (gênero), até então presentes na Internet, que utilizam dados sociais para aumentar sua capacidade de relacionamento, interação, persuasão e influência sobre atores humanos. Isso porque as pessoas deixam cada vez mais dados em rede, que são consumidos por máquinas, as quais são ativas não apenas para auxiliarem humanos, mas também para entender o perfil dos indivíduos que povoam esse ambiente.

Para Thurler (2015, p. 6), as atividades dessas máquinas ultrapassam a perspectiva de compartilhar informações, observando que:

Podem contribuir para executar tarefas relacionadas à recuperação de informações, criação e manutenção de uma base de conhecimento / *wikis*, sistema de recomendação (pessoas e/ou conteúdo), respostas a dúvidas mais frequentes, estímulo ao compartilhamento de conhecimento e incentivo ao reconhecimento e valorização dos profissionais, dentre outros (THURLER, 2015, p. 6).

Desta forma, constata-se que, ao contrário do que se imaginava até então, a Internet é um universo povoado por agentes humanos e não humanos. É possível verificar a grande incidência de agentes autônomos (autômatos) que deixaram de ser mero suporte tecnológico para o processamento informacional, inteligente, assumindo um papel de protagonista na coleta, tratamento, análise e disseminação de dados na rede mundial de computadores.

Floridi (2014) traz uma perspectiva quanto ao crescimento da população mundial e de dispositivos conectados chegando à conclusão que, entre o período de 2001 a 2020, a população humana passará de seis para oito milhões, enquanto no mesmo período o número de dispositivos conectados à rede (Internet das coisas) saltará de praticamente zero para cinquenta milhões de sistemas de *software*. Desta forma, sugere que o processamento de informação ultrapassaria a relação humano-máquina, fazendo surgir um cenário de troca de informações entre máquina-máquina. Ou seja, se a perspectiva de Floridi se concretizar, em 2020 cada ser humano teria para si uma média de seis dispositivos conectados à Internet.

O que se verifica, portanto, é um protagonismo das máquinas nas redes, que se configura pelo “pouco controle dos seres humanos sobre a Internet com a potencialização do uso das informações por parte de máquinas” (SANTANA; LIMA, 2018, 314). Surge, assim, uma tendência à diminuição do controle por parte dos humanos no que diz respeito ao conteúdo veiculado na Internet e um aumento do número de máquinas capazes de processar essa informação. Entretanto, muitas dessas entidades são “silenciosas” e por vezes passam despercebidas por usuários humanos.

A partir desse diagnóstico é possível notar que a interação máquina-homem consolidou-se com a *Web 3.0* e, a partir da evolução da *Web Semântica*, surge um ambiente propício para o desenvolvimento dessas entidades (máquinas) que antes encontravam um ambiente hostil para localizar e indexar informações, em razão da forma precária como essa era disponibilizada em rede. Neste sentido, Santana e Lima (2018, p. 311) pontuam que “as ferramentas de busca passaram a ser mais eficientes, a informação ficou mais fácil de ser encontrada pelos seres humanos e foram minimizados os problemas relativos à falta de informação”.

De fato, encontrar informações dispostas na rede torna-se uma tarefa fácil, mas converte-se ainda mais simples para as máquinas que passam a percorrer a Internet com maior facilidade, especializando-se na coleta de dados disponibilizados em redes que resultaram no aparecimento de agentes inteligentes capazes de atuar nas redes sociais virtuais, inserindo-se no fluxo informacional que prevalece nesses espaços.

Assim, deve ser considerado que a informação, quando tratada a partir da perspectiva do usuário, surge como instrumento utilizado nas relações sociais com a finalidade de influenciar opiniões, redirecionar esforços, mudar objetivos, decisões ou até mesmo ações. Partindo desse pressuposto, não se pode desconsiderar que esta manipulação também consiga ser realizada por agentes não-humanos.

Não é incomum que a influência em rede se dê pelo uso de algum mecanismo de desinformação. Essa última deve ser observada como um fenômeno natural, social, que pode ter origem numa informação falsa (ou não), com a finalidade de auxiliar (ou prejudicar) outrem, e que vem ocorrendo regularmente no cotidiano das pessoas na Internet. Entretanto, pouco se discute sobre as investidas de máquinas nesse processo.

Sobre esse cenário é possível notar um aumento exponencial de episódios que desencadearam processos de desinformação, envolvendo a ação de atores não humanos que potencializam a disseminação de informações falsas/semifalsas/distorcidas ou enviesadas, atingindo em certo grau a sociedade, principalmente no contexto político, com o intuito de beneficiar ou prejudicar a imagem de determinada personalidade pública.

Para Bezerra, Capurro e Schneider (2017), no ambiente *online* um volume imensurável de desinformação vem sendo propiciado por máquinas, tornando-se um problema para plataformas como o *Facebook* e o *Twitter*, por exemplo. Apesar de os autores apontarem as duas principais redes sociais utilizadas atualmente no mundo, não se pode ignorar o avanço dessas investidas no *Whatsapp*, que se tornou instrumento de desinformação e campanha política nas eleições de 2018, no Brasil.

A manipulação de informação por essas entidades que acabam por desencadear processos de desinformação em grande escala é observada por Santana e Lima (2018) que afirmam:

Os robôs já manipulam mais informação na Internet do que os seres humanos. Isso significa que já existe uma estruturação mínima nas informações contidas na rede e um grau de sociabilidade dessas máquinas que sugerem que a tendência é que o tráfego na web seja cada vez maior e com a maior participação de robôs (SANTANA; LIMA, 2018, p. 313).

Na visão Haustein (2016) essas máquinas podem espalhar grande volume de notícias benevolentes ou informações inverídicas, descontextualizadas para influenciar opiniões, por isto, parte desses agentes buscam a confiança de usuários humanos nas redes para depois influenciar opiniões.

Assim, Boshmaf *et al.* (2012, p. 12) citado por Haustein (2016, p. 233) afirmam que é comum para um *bot* “infiltrar-se em comunidades *on-line*, construir confiança sobre tempo e, em seguida, enviar mensagens personalizadas para obter informações, influenciar opiniões e chamar a atenção”.

Desta feita, é possível relacionar esse acontecimento envolvendo a manipulação de informações por partes dos robôs, isso se deve ao fato das máquinas terem mais acesso à informação na *Web 3.0*, conforme apontado anteriormente, o que possibilita o aprimoramento na sua forma de trabalhar.

Em Freitas, Benevenuto e Veloso (2014) é possível constatar que diante do poder de influência desses agentes (não humanos) nas redes sociais, estes são usados durante o período de campanha eleitoral por políticos para aumentar, artificialmente, sua popularidade nas redes sociais, número de seguidores, além de alterar os *Trending topics* e propagar ações publicitárias e políticas, passando a impressão de que se trata de um movimento social popular e espontâneo. Fenômenos como esses passam a ser destaques numa realidade cada vez mais próxima e corriqueira em diversos países do mundo, a exemplo disso, tem sido identificada a ação dessas entidades em campanhas eleitorais realizadas nos Estados Unidos (2012), Itália (2012), Ucrânia (2013), Brasil (2014), Equador (2014), México (2014), Rússia (2014), Turquia (2014), Venezuela (2014), Brexit (2016), China x Tibet (2016) e, novamente, Brasil (2018).

Fato é que as repercussões que decorrem dessas ações já são observadas no mundo real, pondo em risco a integridade das pessoas, a soberania dos países, os processos democráticos e o direito à informação. Dessa forma, a problemática passa a receber atenção de acadêmicos, profissionais da mídia e até mesmo de pessoas comuns, devido ao impacto que esse fenômeno tem trazido à sociedade como um todo.

Ressalta-se que esses acontecimentos são potencializados quando ocasionados por agentes não humanos, posto que essas entidades contam com grande capacidade de compartilhamento de conteúdo nas redes sociais, agindo de maneira automatizada na disseminação das informações. Ora, diante disso, surge a necessidade de entender como vem se estabelecendo esse fenômeno, onde a informação é distorcida e utilizada por máquinas, para que seja possível a sua compreensão e a partir disso sejam estabelecidos planos de contingenciamento, prevenção e conscientização daqueles que estão imersos na Internet.

Entretanto, nota-se que, apesar da desinformação provocada ou ampliada por máquinas sociais na Internet ter se tornado um fato recorrente, o tema ainda carece de discussão e análise por meio da perspectiva da Ciência da Informação (CI), tendo em vista a necessidade de preocupar-se com os desdobramentos sociais que envolvem a (des) informação nesse ambiente, mas que, até agora, vem sendo postergada e delegada para outras áreas distintas que não tem observado a problemática por esse aspecto.

Além de um aprofundamento teórico na discussão, pretende-se investigar casos em que foram verificadas a participação de agentes de *software* em processos de desinformação para responder aos seguintes questionamentos: **Qual é a participação das máquinas sociais no processo de desinformação? Essas são capazes de influenciar o que acontece fora do mundo virtual?**

Assim, faz-se importante determinar o papel de Máquinas Sociais atuantes na Internet concebidos especificamente para influenciar opiniões e induzir comportamentos de usuários (humanos) a partir da desinformação. Então, compreender a relação entre agentes não humanos → pessoas, exatamente no sentido máquina → seres humanos, na Internet, se mostra importante para determinar o poder que pode ser exercido neste tipo de relação.

É deste contexto que vem a motivação da pesquisa, a qual tem como objetivo geral analisar a participação de máquinas sociais (*software*) no processo de desinformação que ocorre na Internet. Uma vez que, mesmo a Ciência da Informação já possui seu olhar para o objeto Desinformação, ainda se faz necessário desenvolver esta visão na Internet e com atores não humanos.

Para isso foram elencados os seguintes objetivos específicos que auxiliam na compreensão do cenário supracitado:

- (a) Compreender o fenômeno da desinformação;
- (b) Identificar e Classificar os atores não humanos, a partir de sua tipologia, e evidenciar a escolha pelas máquinas sociais;
- (c) Observar as possibilidades teóricas da relação entre Desinformação x Máquinas Sociais;
- (d) Compreender a participação das máquinas nos processos de desinformação a partir do estudo de caso;
- (e) Sugerir as repercussões dos fenômenos de desinformação cuja participação das máquinas foi relevante.

Ademais, cumpre destacar que os termos empregados no referencial teórico para referir-se as entidades não humanas que participam dos processos de desinformação deflagrados na rede estão conforme foram empregados originalmente por seus autores; sendo estes pontuados individualmente quanto as suas características no capítulo 3 (entidades Máquinas Sociais e entidades não biológicas).

Desta feita, não custa iniciar o presente instrumento explanando as diversas facetas e manifestações da desinformação que resultam no surgimento de movimentos cada vez mais presentes na Internet, como o vigilantismo, checagem de fatos, filtros bolhas e tecnocensura, que serão detalhados a seguir.

2 DESINFORMAÇÃO

Neste capítulo será tratada a desinformação partindo do seu conceito e as principais visões do assunto na Ciência da Informação. Em seguida, será traçado um paralelo entre informação e desinformação, baseado na visão de Buckland (1991), posto a antítese que se observa entre ambas. Posterior a isso serão apresentadas manifestações da desinformação e como estas vêm se materializando na sociedade e principalmente na Internet. Por fim, serão apresentados alguns dos movimentos que surgiram em virtude da desinformação.

Iniciando por Pinheiro e Brito (2014), estes afirmam que desinformação deve ser compreendida como um fenômeno negativo da informação, corrobora assim, o entendimento firmado anteriormente por Capurro (2003, p. 1, tradução nossa) o qual afirma que “no domínio da ciência da informação existe a preocupação com a forma negativa, a desinformação e seus derivados, mentiras, má interpretação, ilusão, erro, decepção”.

Segundo Watson *et al.* (1990) o termo desinformação tem origem na palavra russa “*дезинформация*”, usada no sentido de falsificar informações, de maneira intencional, particularmente lançadas por um governo para atingir outro ou dirigido para um determinado público. Serviços de inteligência definem desinformação de maneira similar, o termo é usado em operações de inteligência secretas e dificilmente descobertas para descrever evidências falsas e deturpadas utilizadas para desacreditar o oponente. No mesmo sentido Zattar (2017, p. 288) afirma que “a noção de desinformação surge no contexto das práticas de guerra”, acompanhando este raciocínio Volkoff (2004) e Fallis (2015) sugerem a mesma gênese para o termo.

Stahl (2006) divide a desinformação em duas facetas: (i) *Disinformation* que é o ato de disseminar informações incorretas de maneira intencional e (ii) *misinformation* que se refere à informações incorretas ou enganosas, disseminadas sem a total compreensão de que estas sejam falsas; a diferença dentre estes termos reside, portanto, na intenção de quem disponibiliza a informação.

A Desinformação tem recebido a atenção de acadêmicos, profissionais da mídia e até mesmo de pessoas comuns, devido ao impacto negativo que este fenômeno tem

trazido a sociedade. Questões ligadas a política, causas sociais, religiosas e nas trivialidades dos tabloides sofrem algum tipo de influência da desinformação que leva pessoas a assumir/acreditar em fatos inverídicos. A questão da desinformação não passa apenas pela criação e disseminação de falso conteúdo nem pela momentânea exposição à mentira em si, mas sim pelo processo desencadeado pela memória humana em acomodar aquela visão de mundo como verdadeira, levando o indivíduo a crer que tal acontecimento/opinião é um fato verdadeiro.

Loftus (2005) aponta que o processo de desinformação é bastante complexo e envolve diversas questões não só psicológicas, mas também fisiológicas, comportamentais e até mesmo de crenças. A autora trata a questão como algo que vai além do conteúdo inverídico em si, como objeto e seus processos de propagação, mas sim conforme a transformação mental do indivíduo e do assentamento daquela mentira como permanente. Segundo ela, a desinformação também é um processo de criação, não de memórias falsas, mas sim de memórias reais que assumem o *status quo* da mentira como fato.

É no turbilhão da explosão informacional que também viraliza a questão da desinformação. Com a popularização da Internet, e em especial das redes sociais, toda a sociedade encontrou um espaço de fácil e livre disseminação de opiniões e, conseqüentemente, apropriou-se de mais um meio para compartilhar conteúdo inverídico. Pessoas movidas pelos mais diversos interesses promovem, em rede, uma série de novos comportamentos informacionais que são amplamente divulgados em veículos especializados em tecnologia, noticiados nos meios de imprensa e periódicos acadêmicos. Novas maneiras de se relacionar com outros indivíduos, novas formas de se comunicar, novos regimes de trabalho e até mesmo novos hábitos de compra se consolidaram com a combinação da Internet e *Smartphones*. Este cenário contemporâneo se torna prolífico para a difusão de conteúdo inverídico tornando a Internet um ambiente rico para estudos relacionados a desinformação.

De acordo com Matheus (2005, p. 156), “a Ciência da Informação deve considerar a informação e a desinformação como objetos complementares”. Schrader (1986) critica o excesso de zelo da CI pela amplitude de definições dada ao objeto informação, uma vez que, por outro lado, quase não havia referência à forma negativa, chamada de

desinformação, e seus derivados que seriam "mentiras, propaganda, deturpação, rumores, alucinação, ilusão, erro, ocultação, distorção, embelezamento, insinuações e enganos" (SCHRADER, 1986, p. 179, tradução nossa). Segundo o autor essa visão positiva sobre a informação resulta em um "modelo ingênuo do homem informacional, que às vezes toma a forma de homem tomador de decisões ou de homem da incerteza" (SCHRADER, 1986, p. 179, tradução nossa).

Capurro (1992, p. 85, tradução nossa) faz o seguinte questionamento "para que serve informação?", que leva a outro questionamento "Qual é a finalidade da Ciência da Informação?" (CAPURRO, 1992, p. 85, tradução nossa). Para o autor, a CI é uma disciplina "Hermenêutica-Retórica" que investiga "dimensões pragmáticas e contextuais" dentro do contexto em que conhecimento é compartilhado de forma positiva como informação e negativa como desinformação, em particular a partir de formas técnicas de comunicação.

Capurro (2003) também ressalta a quase inexistência de pesquisas sobre esta forma negativa. Para o autor, existe o entrelaçamento entre informação e a sua antítese como "dimensões existenciais" e que ambas são pseudônimos para uma série de experiências, tais como, aquelas listadas por Schrader (1986), mentiras, propaganda e enganos, para o lado negativo e que foram contrapostas por ele no lado positivo, tais como, "retidão, franqueza, clareza, utilidade e veracidade". Assim, o autor conclui que a CI é uma ciência tanto para a informação quanto para a desinformação.

Tudjman e Mikelic (2003) questionam se a desinformação realmente existe como um assunto da prática de documentação da informação. Segundo os autores, a CI não pode negligenciar o fato de que a troca de informações é uma relação social cujos participantes muitas vezes têm interesses opostos. Estes atores se esforçam para mudar as intenções, objetivos, decisões e ações – não apenas dos envolvidos diretos, mas também, de outros participantes indiretos com os quais possam estar em algum tipo de conexão. Ou seja, para o autor, a desinformação é uma manifestação natural na relação entre seres humanos e não um fenômeno voltado ao engodo. É por isso que a Ciência da Informação deve começar a explorar a relevância da informação do ponto de vista do usuário, bem como do ponto de vista dos remetentes e de todos os outros participantes da interação.

Floridi (1996) aponta que a desinformação surge sempre que o processo de informação é defeituoso. O autor dá a entender que todo o contexto da Ciência da Informação existe para que o processo de informar ocorra de modo eficaz. Mas, e quando esse processo falha? Essa desinformação, segundo ele, precisa também ser investigada.

Zattar (2017) aponta que já existem trabalhos relacionados à desinformação no campo de estudos da informação. A autora exemplifica com a “*Resolution on disinformation, media manipulation, and the destruction of public information*” (AMERICAN LIBRARY ASSOCIATION, 2005) e a “*Resolution on access to accurate information*” (AMERICAN LIBRARY ASSOCIATION, 2017) que são esforços da comunidade da CI no sentido de prover o conhecimento de fontes de informação para o reconhecimento da desinformação.

Apesar de verificar certa semelhança e consenso quanto à origem do termo desinformação que surge no contexto das guerras e o seu sentido negativo face a informação, a desinformação se mostra plural, da mesma forma que o conceito informação, podendo se referir a diferentes aspectos informacionais. Sendo assim, faz-se necessário estabelecer um paralelo entre informação e desinformação, ressaltando a antítese que há entre esses dois objetos, por isso, é necessário entender que estes são indissociáveis, mas que apresentam características opostas quanto aos seus desdobramentos na sociedade.

Fallis (2015) sugere que, ao tratar de desinformação, podem ser inferidas três observações: (i) **desinformação é informação**; (ii) **desinformação é uma informação enganosa** e (iii) **desinformação não é uma informação acidentalmente enganosa**. Esta classificação mostra três aspectos sobre desinformação e que ao mesmo tempo aproximam esta da informação, possibilitando identificar que estes são opostos.

Em **primeiro lugar**, o autor afirma que “desinformação é um tipo de informação a depender em qual análise da informação nós escolhemos adotar” (FALLIS, 2015, p. 404, tradução nossa). Desta forma, trata de resgatar a ideia de que não há uma distinção clara, de lados ou polarizações, entre informação e desinformação. Ou seja, existe uma zona cinza que envolvem as meias verdades, ou questões de contexto que podem fazer com que a fronteira entre estas duas questões se tornem tênue ou nebulosa. Para Fallis

(2015, p. 405, tradução nossa) a utilização do termo informação pode se dar “para se referir ao conteúdo representacional que é falso, bem como ao conteúdo representacional que é verdadeiro”.

Essa visão aproxima-se de Scarantino e Piccinini (2010) que previram a possibilidade do termo informação ser utilizado tanto para se referir aquilo que é verídico, quanto para aquilo que não o é. Entretanto, afasta-se da ideia pregada por Floridi (2011) que enxerga o elemento informação como algo que se refere apenas à verdade. Para o autor, a desinformação ocorre quando existem falhas no processo de informação, existindo, assim, uma visão dual e antagônica entre os termos (FLORIDI, 1996).

No **segundo ponto**, ao aproximar desinformação à informação enganosa, Fallis (2015) não visa tratar da presença ou ausência de intencionalidade quanto à disseminação deste conteúdo, mas esclarece que “é a informação que provavelmente cria falsas crenças” (FALLIS, 2015, p. 406). Ou seja, a informação é vista como um meio que viabiliza a concepção da ação de desinformar. Destaca o autor que, mesmo não atingindo o seu alvo, a desinformação ainda persistirá mudando apenas o seu grau de consequências no plano material. Sendo assim, trata da abrangência do que pode ser desinformação se referindo a qualquer problema ligado à informação, seja este intencional ou não. A ideia é abarcar o termo *misinformation* (informação não intencional) como um todo e mostrar que podem existir outras questões ligadas à desinformação que não a intencionalidade.

Por último, o autor retrata a intencionalidade que reside na propagação da informação enganosa como característica da desinformação, aproximando-se da *disinformation* (informação intencional) já tratada anteriormente, não se confundindo, portanto, com erros, enganos ou sátiras.

Muito próximas da caracterização delineada por Fallis (2015) estão as considerações de Pinheiro e Brito (2014) que visualizam a desinformação por três prismas, a saber: (i) **Ausência de Informação**; (ii) **Informação Manipulada** e (iii) **Engano Proposital**.

A **Ausência de Informação** é relacionada, indiretamente, por Aquino (2007) ao conceito de desinformação quando se constata que nas interações que os indivíduos mantêm com o mundo e com os outros, eles estão, quase sempre, em estado de

desconhecimento total ou possuem pouca informação sobre um assunto. Neste cenário, a forma de pensar da sociedade seria separada entre informados e desinformados, sendo os últimos um conjunto de “excluídos do acesso a bens informacionais” e tecnológicos.

Segundo Pinheiro e Brito (2014), a ausência de informação é a situação mais amplamente associada ao termo desinformação. Eles até exemplificam com a definição de desinformação do dicionário Michaelis, como sendo o “estado de uma pessoa ou grupo de pessoas não informadas ou mal informadas a respeito de determinada coisa”. Esta situação sugere uma precariedade informacional do sujeito devido à sua própria ignorância. Assim, desinformação resulta em ausência de cultura ou de competência informacional tornando impossível para o usuário localizar por si só a informação desejada.

Compreende-se, portanto, que os autores não tratam da completa ausência de informação dos indivíduos, pois consideram, também, entre os sujeitos submetidos à desinformação aqueles que possuem acesso à subinformação, seja esta parcial ou incompleta. Castro e Ribeiro (1997) sugerem o aparecimento da sociedade da desinformação, do analfabetismo tecnológico e dos excluídos digitais com diferentes níveis de acesso aos bens culturais, sendo possível aferir a diferença entre os agentes envolvidos na sociedade da informação com acesso aos recursos tecnológicos e bens culturais, mas que ainda assim estão sujeitos a processos de desinformação causados pelo excesso de conteúdo, sua baixa qualidade e outros fatores.

Ao tratar da **Informação Manipulada**, é considerada a visão de Demo (2000), a qual aponta que a desinformação pode ser o objetivo principal de determinados grupos. Não se tratando apenas de apresentar uma massa de informação, de tal modo que, seja impossível o manejo dela, mas sobretudo de usá-la para o propósito oposto, no “sentido mais preciso de cultivo da ignorância”. Para o autor, nem a Internet conseguiu modificar este modelo desinformador, em que “a habilidade inovadora do conhecimento não é menor quando motivada por projetos colonizadores” (DEMO, 2000, p. 39). Na verdade, o autor alerta que os modelos informacionais como os da sociedade da informação existem em um papel oposto ao apregoado, em que as benesses estariam mescladas com a informação cujo propósito é de desinformar. Assim, seria “sempre possível, pois,

usar o melhor conhecimento para construir o mais refinado processo de imbecilização. Desinformar será, portanto, parte fundamental do processo de informação” (DEMO, 2000, p. 39).

Para Demo (2000), redes como a Internet aceleram a propagação da informação de baixo valor cultural ou utilidade, ao que denomina desinformação. Percebe-se que esta visão identifica a “sociedade da desinformação” como algo em desenvolvimento, associado a este “novo tipo de sociedade”. Assim, a desinformação consiste em um grande conjunto de informações disponibilizadas cotidianamente, mas que não oferece ao indivíduo o conhecimento especializado para participar do processo político e tomar decisões relativas ao progresso de sua própria vida e de seus semelhantes. Propositadamente, estas informações de baixo valor seriam fruto de um projeto de dominação, em que tanto as redes digitais, quanto veículos de comunicação tradicionais estão engajados.

Corroborando, neste sentido, Pinheiro e Brito (2014) observam que a informação manipulada retrata a situação presente sobre a desinformação e está relacionada ao fornecimento de produtos informacionais de baixo nível cultural, resultando na “imbecilização” de determinados setores sociais. Nesta concepção, os setores da elite desinformam amplamente com o intuito de se perpetuarem no poder, atingindo mais facilmente seus próprios interesses.

Por fim, ao tratar do **engano proposital**, se remonta à ideia do termo original em russo “дезинформация”, que como a palavra desinformação foi criada para tal função. Pinheiro e Brito (2014) apontam que o caráter intencional da ação, a aspiração de enganar outrem é parte determinante do conceito. Segundo Carvalho (2001), a maior parte das pessoas instruídas não sabe o que é desinformação, imagina que se trate apenas informação falsa para fins gerais de propaganda, ignorando por completo que o principal foco são ações perfeitamente calculadas para uma finalidade específica, e que em grande parte não é influenciar as multidões, mas atingir alvos muito determinados para induzi-los à tomada de decisões prejudiciais a seus próprios interesses e aos de seu país.

Nesta visão, a desinformação, decepção e operações psicológicas não são distinguíveis umas das outras mudando apenas a palavra “propaganda” para o “alto

nível”. Destarte, apontam Pinheiro e Brito (2014) que enquanto “a desinformação na propaganda lida apenas com dados políticos ao alcance do povo, a desinformação de alto nível falseia informações especializadas e técnicas de relevância incomparavelmente maior”.

Considerando Fallis (2015) e Pinheiro e Brito (2014), é possível notar que a informação e desinformação apresentam efeitos antagônicos no que diz respeito à sua aplicabilidade e consequências quando postas à disposição da sociedade. Verifica-se, portanto, o caráter positivo de uma e o negativo de outra, entretanto, estão dispostas de maneira indissociável, razão pela qual Demo (2000) pontua que desinformar pode ser o projeto principal da informação, isso porque, se utilizada para o seu propósito oposto, servirá para cultivar a ignorância.

Para o autor, a desinformação faz parte da informação, assim como “a sombra faz parte da luz”. Defende, portanto, que trata-se do mesmo fenômeno, com sinais invertidos e apresenta como exemplo que estudos sobre conhecimento apontam para esta característica ambivalente, a começar pela ideia de “conhecimento proibido”.

Desta forma, enquanto a informação apresenta um caráter disruptivo permitindo ao sujeito o poder de questionamento, construção e desconstrução da realidade, a desinformação surge justamente quando essa capacidade não está a serviço de propósitos lícitos, transparentes e coletivos, razão pela qual “é sempre possível, pois, usar o melhor conhecimento para construir o mais refinado processo de imbecilização” (DEMO, 2000, p. 38).

Outro fator que merece destaque é o caráter manipulativo da informação no sentido negativo em detrimento da informação no sentido positivo. Para Demo (2000), o que expõe a sociedade a níveis de desinformação é (i) Informação residual; (ii) Informação de classe superior e inferior, cuja variação está em função como regra do poder aquisitivo de cada um; (iii) Informação imbecilizante, seja por conteúdo medíocre/fútil, seja por distorções/ clamor exacerbado; (iv) A mídia, que não exerce sua função pública, visando apenas interesses comerciais; (v) Os novos meios de comunicação, que ainda servem em grande parte a elite.

Observa-se também uma disparidade entre a qualidade do conteúdo apresentado na informação e na desinformação, posto a ausência de requisitos como “eficácia,

precisão, completude ou completeza, consistência, relevância, autoridade cognitiva e credibilidade” (CALAZANS, 2008, p. 30) desta em relação àquela. Neste sentido, enquanto a informação de qualidade é adequada ao uso e tem valor para o usuário (SARACEVIC, 1999), na desinformação há ausência desses requisitos justamente por esta não suprir adequadamente as necessidades do usuário que busca pelo conteúdo.

Fallis (2015) reconhece a impossibilidade de uma análise conceitual perfeita para o termo. Para ele,

A desinformação pode ser um “conceito de protótipo”; isto é, pode haver simplesmente instâncias prototípicas de desinformação, com coisas diferentes ficando mais próximas ou mais distantes desses protótipos. Mas mesmo se é **impossível encontrar uma análise que capture perfeitamente o conceito de desinformação** (isto é, uma que seja “prova de contra exemplo”), ainda podemos nos esforçar para chegar o mais perto possível (FALLIS, 2015, p. 415-416, tradução nossa).

Se ainda uma imprecisão sobre a desinformação aparentar ter múltiplas visões, para a informação isso é bastante claro. Por exemplo, Buckland (1991) acreditava que estabelecer um único significado para a informação resultaria na incompreensão de outros usos do termo, limitando as explicações quanto a sua aplicabilidade. Por esta razão, baseando-se no que previa o *Oxford English Dictionary* de 1989, ousou, apresentando as facetas pelas quais a informação poderia se apresentar, neste sentido, pontuar que a informação poderia ser vista como coisa, conhecimento ou processo.

Para o autor, a **informação como processo** está vinculada ao ato de informar ou comunicar alguém sobre um fato ou ocorrência, modificando sua carga cognitiva. A **informação como conhecimento** é “usada para denotar aquilo que é percebido em informação-como-processo” (BUCKLAND, 1991, p. 351, tradução nossa), desta forma, defende que se trata da informação que reduz ou aumenta a incerteza. Por fim, a **informação como coisa**, noção amplamente empregada na CI e que estaria relacionada à tangibilidade dos objetos como dados e documentos, meio pelo qual se transmite (ou materializa-se) o conhecimento, utilizada para fins de instrumentalização. Desta forma, como bem pontua:

A distinção entre intangíveis (conhecimento e informação-como-conhecimento) e tangíveis (informação-como-coisa) é central para o que se segue. Se você pode tocá-lo ou medi-lo diretamente, isso não é conhecimento, mas deve ser

algo físico, possivelmente informação-como-coisa (BUCKLAND, 1991, p. 351, tradução nossa).

Assim, ao contextualizar sua teoria, Buckland (1991, p. 352, tradução nossa) trata da “intenção para que os usuários se tornem informados (informação-como-processo) e que haja uma transmissão de conhecimento (informação-como-conhecimento)”, destacando que, “o meio fornecido, o que é manuseado e operado, o que é armazenado e recuperado, é informação física (informação-como-coisa)” (BUCKLAND, 1991, p. 352, tradução nossa).

Ora, se a informação apresenta facetas, é possível identificar que a desinformação também tem as suas e, desta forma, aproveitando-se da proximidade de ideias apresentadas por Buckland (1991) aqui se propõe delinear que a desinformação pode ser enxergada como **coisa, processo e propósito**, conforme será discorrido a seguir. Não se tem aqui a intenção de fechar ou afunilar as dimensões sobre as quais a desinformação pode ser observada, muito pelo contrário, aqui se tenta abrir este leque propondo a ampliação deste debate.

A escolha destas três categorias se deu pela proximidade com as ideias de Buckland (coisa e processo) e pela visão de outros autores da CI que tratam de aspectos de intencionalidade, objetivo e propósito de desinformar. Essa visão, por exemplo, é compartilhada por Pinheiro e Brito (2014) que visualizam a desinformação por três prismas já mencionados: (i) Ausência de Informação (coisa); (ii) Informação Manipulada (processo) e (iii) Engano Proposital (propósito).

2.1 Desinformação como Coisa, Processo e Propósito

A informação assume em Buckland (1991) facetas que permitem sua aplicabilidade e estudo a partir de diferentes pontos de vistas complementares. Aqui se assume que o mesmo também é válido para a desinformação, ou seja, a sua aplicabilidade e uso definirá qual a melhor abordagem para analisar a desinformação. Sendo assim, optou-se neste trabalho por estruturar a desinformação enquanto coisa, processo e propósito, entretanto, cabe ressaltar que a configuração desta como uma ou

outra faceta não exclui sua aplicabilidade quanto aos demais sentidos, visto que a mesma desinformação poderá assumir mais de uma faceta.

2.1.1 Desinformação como Coisa

Seguindo, esta análise pode ser iniciada partindo da **desinformação como coisa**. Antes de tudo, ressalta-se que não se mostra adequado tratar esta faceta da desinformação de forma estritamente oposta à da informação, pois, se em parte Buckland relaciona esta faceta ao suporte, a desinformação se utiliza destes mesmos meios para prosperar. Então, não é estritamente a materialidade nem os meios de transmissão que se deve aqui observar e sim, inicialmente, a informação enquanto objeto que existe, persiste, e que pode possuir falhas.

Floridi (2012) sugere que a desinformação pode ser utilizada para se referir às “informações semânticas falsas”. Objetos que possuem falhas em sua estrutura e que, devido à sua imperfeição, passam a desinformar mais do que informar. Fallis (2009) relaciona diretamente a desinformação com a qualidade da informação em uma relação proporcionalmente inversa. Quanto maior a qualidade da informação, menor as chances de existirem desinformação naquele contexto. Desse modo, o autor extrapola questões meramente de significado (semântica) e passa a questionar relações mais abrangentes entre os dois elementos em questão, tais como contexto, temporalidade e uso adequado da informação. Assim, pode-se inferir que a desinformação também pode ser, dentre tantas outras possibilidades, a informação com baixa qualidade. Seguindo estas ideias, pode-se dizer que esta vertente apresenta a desinformação como **falha no objeto informação**.

Floridi (2012) apresenta a ideia de que a desinformação pode ocorrer devido à problemas na qualidade do objeto informação. Fox (1983, p. 75, tradução nossa) afirma que “quando o conteúdo semântico é falso, este é um caso de desinformação”. Esta afirmação é corroborada por Floridi (2012) que posiciona as mentiras, engodos e falsidades no contexto da desinformação como o repasse de informação inverídica. Este olhar para a desinformação é aquele onde se encaixam as notícias falsas (*Fake News*)

e é, também, o que relaciona diretamente a desinformação à informação de baixa qualidade.

Entretanto, Fallis (2009) ultrapassa esta visão ao apresentar o conceito de que a falha no objeto semântico e no objeto informação nem sempre se dá por questões de correteude e precisão, e sim, por questões também contextuais que podem envolver temporalidade, credibilidade e origem da fonte de informação. O autor afirma que, diferentemente de uma mentira, a desinformação não precisa de uma declaração formal, mas que esta pode partir de uma verdade, que possui detalhes omitidos, de maneira não proposital, inclusive, que leve à desinformação.

Floridi e Illari (2014) afirmam que olhar a desinformação apenas a partir de mentiras, pode afunilar a visão do fenômeno, de modo a tornar a sua análise simplória, rasa e que não abrange todas as situações em que ela se manifesta. Os autores criam uma categoria chamada de “*true disinformation*” que se trata dos enganos que ocorrem a partir de verdades incompletas, mas que ainda assim são verdades com lacunas que podem ser preenchidas de diferentes formas pelo receptor. Um exemplo apresentado pelos autores é um hipotético vilão que, com propósitos maléficis, pergunta a alguém sobre o paradeiro de um amigo, e esta pessoa pode responder algo genérico, embora verdadeiro, do tipo: “Ele adora jantar em São Paulo”. Mas que São Paulo? O estado? A cidade? O clube de futebol? Uma rua em algum outro lugar? Uma igreja? Um restaurante? Enfim, se percebe que a resposta é imprecisa, embora não seja, necessariamente, mentirosa.

Fallis (2015) classifica a situação acima como uma meia verdade, uma vez que somente parte da informação possui qualidade adequada, e outra parte possui qualidade baixa, sendo o suficiente para se tornar dúbia e imprecisa. O interessante da visão de Fallis é que a informação não é uma entidade indivisível, una e homogênea. Para ele, uma determinada informação, considerando-a como objeto, pode apresentar diferentes níveis de qualidade em toda a sua extensão. Saracevic (1999) afirma, de forma abrangente, que a informação de qualidade reconhecida é adequada ao uso e tem seu valor percebido pelo usuário. Quando ocorre a desinformação, são justamente esses requisitos de qualidade que se tornam ausentes e assim a informação passa a não atender adequadamente as necessidades do usuário.

Outra forma de se observar a desinformação como coisa é quando existe a **falha do registro da informação em seu suporte**. Um exemplo específico de problemas na qualidade da informação, devido a falhas no suporte, é trazido por Floridi e Illari (2014). Os autores afirmam que se a desinformação é um tipo de informação, ela precisa possuir, de alguma forma, um modo de ser representada e percebida pelos interagentes, e que o problema pode estar no meio e não na informação. E exemplificam que a desinformação pode se originar de informações sensoriais, que são propositadamente criadas, tais como: falsos mapas e montagens em fotografias, ou mesmo acidentais, ou ainda, como um sinal sonoro com ruído, ou tentam obter informações de um papel borrado pela chuva.

Arif (2018) faz uma relação desta forma de desinformação com a preocupação e a manutenção da informação em suportes devidamente preservados. Nesse sentido, o autor relaciona o conceito de desinformação ao esquecimento (ou perda da memória) ao apontar que informações perdidas pela destruição, obsolescência e desuso de uma determinada tecnologia, traz um empobrecimento da compreensão de determinado assunto em um grau, maior ou menor, de desinformação. A visão de Arif é que o esquecimento e o descaso com a preservação das informações são fatores que contribuem para a desinformação.

Benkler, Faris, Roberts e Bourassa (2018) também observam que a desinformação, a partir da concepção de problemas na qualidade da informação, pode se dar pelo meio de transmissão, principalmente quando estes estão se tornando cada vez menos neutros. Assim, é possível elencar outro tipo de desinformação como coisa que é a **falha na transmissão da informação**. Os autores sugerem que a desinformação é um problema fundamentalmente político e social, e não tecnológico, entretanto, a forma como a Internet transmite informação hoje está mais para um “telefone sem fio” do que para uma transmissão direta de informação.

Desse modo, é possível compreender que a desinformação pode ocorrer por mal funcionamento do meio, como por exemplo um telefone celular utilizado em uma área de difícil acesso em que as falas são cortadas durante uma ligação, o que pode levar pessoas a tomarem decisões equivocadas. Todavia, o que os autores chamam a atenção é que existem tantos algoritmos “tratando” a informação disponibilizada para os usuários

que é praticamente impossível que elas cheguem sem algum tipo de incompletude ou viés. Os autores ponderam que a curadoria de conteúdo está migrando para uma curadoria de algoritmos que levam pessoas, de forma cada vez mais intencional, à desinformação.

Hao (2019) aponta que existe uma relação entre desinformação e o uso de inteligência artificial na curadoria de conteúdo, uma vez que os algoritmos têm papel ativo no que é exibido para o usuário. Assim, a autora estabelece uma base para a responsabilização algorítmica para o conteúdo apresentado em meios digitais a partir da desinformação.

Floridi (1996) observa que a desinformação pode ser causada por alguma forma de manuseio incorreto de Informação. Isto quer dizer que esta pode se relacionar à manipulação dos documentos, dados e informações em geral. Segundo o autor, a imperícia em se utilizar os sistemas de informação adequadamente pode levar a um estado de falta de informação, que resulta na incapacidade de se atingir o objetivo esperado.

A última forma, aqui apresentada, de desinformação como coisa ocorre quando esta se dá pela **falha na percepção/compreensão/uso da informação por parte do interagente**. Esta última classificação costuma ser uma consequência decorrente da ocorrência de uma, ou mais, das três falhas listadas anteriormente. Ou seja, falha na informação, e/ou no suporte, e/ou no meio de transmissão resulta na pouca utilidade dela para um usuário. Entretanto, podem acontecer casos em que nenhum dos outros tipos de falha ocorram e, ainda assim, a informação não ser percebida pelo interagente, ou, pode acontecer uma, ou mais, das falhas e o usuário perceber este problema na qualidade da informação e estar ciente de que lida com um possível viés na informação.

Iniciando por Fallis (2015), este observa a desinformação criada por sátiras, piadas e comentários sarcásticos e as posiciona como manifestações da desinformação, quando na verdade, estas não têm a intenção de ser. Então, se compreende muitas vezes que o emissor tem a intenção de se utilizar de um determinado código (sarcasmo, brincadeiras e afins) que não é compreendido pelo receptor.

Uma manifestação de desinformação que está ligada a esta categoria é a pós-verdade. Higgins (2016) sugere que a pós-verdade se trata de um momento onde a

informação que não é conveniente a um determinado interesse passa a ser ignorada em detrimento a um conjunto de crenças que norteiam pessoas, grupos e até mesmo sociedades. Assim, os usuários escolhem, mesmo que de forma inconsciente, não considerar uma informação de maior qualidade por outra de menor, mas que se apresenta de maneira conveniente.

Marshall *et al.* (2015) relacionam a desinformação a problemas de usabilidade em determinados produtos, a partir do momento que o usuário não consegue atingir os seus objetivos devido a não compreensão sobre o funcionamento de determinado dispositivo e que há alguma espécie de comunicação tácita em sua forma de funcionar.

Wardle (2018) apresenta outra manifestação que recai na compreensão dos usuários sobre os dados que é o *Data Flood*. Os usuários estão tão assoberbados de dados, inclusive de qualidade reconhecida, que não conseguem torná-los úteis, revelando uma espécie de deficiência em reconhecer dados relevantes em um grande mar de informação. Usuários humanos normalmente tendem a se confundir mais do que se esclarecer quando estão diante de muitos dados e essa confusão é considerada como desinformação.

Fallis (2015, p. 410) também aponta casos onde a desinformação pode compreender “publicidade enganosa (nos negócios e na política), propaganda do governo, fotografias falsificadas, documentos falsificados, mapas falsos, fraudes na Internet, sites falsos e manipulações de verbetes da *Wikipedia*”. Mas que, pode-se notar que a qualidade da informação foi afetada, sendo, portanto, identificável pelo receptor e que ela não o influenciaria a seguir o comportamento desejado pelo emissor. Então a percepção do usuário sobre aquela informação muda a partir do momento em que ele está ciente da desinformação.

Neste sentido podemos afirmar que, até aqui, a **desinformação enquanto coisa** está ligada a falta de qualidade da informação e que se pronuncia de quatro formas: (i) **falhas ou imperfeições no objeto informação** (ii) **falha do registro da informação em seu suporte**; (iii) **falha na transmissão da informação** e (iv) **falha na percepção/compreensão/uso da informação por parte do interagente**. Essas categorias não são excludentes entre si, e em casos específicos, pode-se observar a ocorrência simultânea de mais de uma dessas formas. Geralmente, os casos de

desinformação como coisa tendem a resultar na última categoria (comportamento inadequado do usuário), mas podem existir situações em que o usuário perceba a falta de qualidade da informação, bem como podem ocorrer questões em que não haja problemas no objeto, no suporte nem no meio que levem o interagente a agir de forma inadequada.

Por fim, deve ser considerada a visão de Aquino (2007) de que, especificamente, “a ausência de informação seja desinformação” e levantar a seguinte questão: esta ausência seria alguma espécie de falha no objeto de informação? Ou seja, seria a falta de informação um fenômeno a ser observado na desinformação como coisa?

Não há dúvidas que essa ausência de informação possa ocorrer por motivos ligados à falhas no devido (ii) registro da informação e em problemas de perdas na (iii) transmissão da informação. Entretanto, deve ser considerado que a forma como cada indivíduo age diante da total ausência de informação é de escolha e responsabilidade dele (processo) ou está mais voltado aos interesses de outrem (propósito) que são atendidos a partir da completa falta de informação. Todavia, ressalta-se novamente que este é um início de reflexão, não se presumindo o esgotamento do debate acerca disto.

2.1.2 Desinformação como Processo

Para iniciar a análise considerando a **desinformação como processo** se deve destacar as duas formas vislumbradas nesta dissertação, ressaltando novamente que não se tem a intenção de limitar a desinformação como processo a estas duas formas, mas sim considerá-las como ponto de partida. A primeira forma (i) é levantada por Floridi (1996) como uma falha no processo de informar, enquanto a segunda (ii) é trazida por Skyrms (2010) quando este se refere à execução de ações sistematizadas (processo) para se atingir o objetivo de desinformar. Deve-se observar com atenção que esta segunda visão, não considera a intenção ou mesmo objetivos em desinformar, mas sim a observação das estratégias elaboradas e utilizadas para desinformar.

Antes de aprofundar as reflexões das duas formas aqui identificadas deve-se salientar que a desinformação como processo é uma visão menos difundida e debatida na academia. Na Ciência da Informação se adota, em sua maioria, a visão que a

desinformação é o aspecto negativo (valor), antagônico (papel) ou até mesmo complementar (objetivos) da informação e assim é natural a compreensão de que observar a desinformação como coisa ou como propósito sejam mais tangíveis e observáveis quando se leva em conta a desinformação partindo do que se conhece como informação.

Entretanto, ao considerar a desinformação como um fenômeno menos dependente da informação, seja no conceito, seja na percepção desta por parte das pessoas, é perceptível que a noção de informação como processo passa a fazer um pouco mais de sentido. Isso tanto é plausível que ao retomar o termo russo original “*дезинформация*” ele está mais ligado ao conjunto de estratégias utilizadas pelo exército do que pelo objeto informação em si.

Nesse sentido, ao considerar a visão de Capurro (2003, p. 1, tradução nossa) ao dizer “[...] desinformação e seus derivados: mentiras, má interpretação, ilusão, erro, decepção”, e a visão de Schrader (1986, p. 179, tradução nossa) que diz “[...] desinformação, e seus derivados que seriam “mentiras, propaganda, deturpação, rumores, alucinação, ilusão, erro, ocultação, distorção, embelezamento, insinuações e enganos”, tem-se uma visão da própria Ciência da Informação sobre este fenômeno. Assim, a desinformação como processo surge como uma visão complementar as outras duas visões aqui elencadas.

Considerando a visão de Floridi (1996), esta pode ser considerada a primeira a circular em meios da CI a observar que a **desinformação é o resultado de um defeito no processo de informar** e que surge “sempre que estão ausentes requisitos como objetividade, integridade e pluralismo, podendo apresentar-se cumulativamente ou não” (FLORIDI, 1996, p. 512, tradução nossa). Desta forma, o autor ressalta que os problemas mencionados podem se referir a uma elaboração errada de fluxos de informação (processo de informar defeituoso) ou que o próprio fluxo de informação não foi seguido da forma planejada (não execução do processo).

O processo defeituoso é aquele que, segundo o próprio autor, diminui ou se elimina os “[...] requisitos de objetividade, integridade e pluralismo” (FLORIDI, 1996, p. 512, tradução nossa). Então, o autor sugere que podem existir processos de baixa

qualidade que resultem na diminuição da utilidade da informação por parte não só do usuário, mas de todo o ecossistema envolvido no contexto daquela informação em si.

Esta visão é complementada por Calazans (2008, p. 30), uma vez que o autor observa entre a qualidade nos processos apresentados para informação e desinformação, posto a ausência de requisitos como “eficácia, precisão, completude ou completude, consistência, relevância, autoridade cognitiva e credibilidade” desta em relação àquela.

Esta maneira de visualizar a desinformação se aproxima daquela vista na desinformação como coisa em que há a dificuldade na percepção/compreensão/uso da informação por parte do interagente, entretanto, aqui a falha está no processo de informar (comunicação) e não nas limitações, de qualquer natureza, do usuário. Enquanto, a desinformação como coisa (percepção/compreensão/uso da informação por parte do interagente) se dá exclusivamente pela inabilidade/inadequação/impossibilidade do usuário compreender a informação, aqui o problema está na forma como foi concebido o processo de informação e dos problemas inerentes deste.

Capurro (1992) separa as questões da desinformação como coisa e da desinformação como processo ao afirmar que dentro do contexto que conhecimento é compartilhado de forma positiva como informação e negativa como desinformação, “em particular a partir de formas técnicas de comunicação” (CAPURRO, 1992, p. 87). Compreende-se, então, que a forma negativa (da informação) também pode se dar pelo processo de comunicação.

Fallis (2015), ao afirmar que a informação é vista como um meio que viabiliza a concepção da ação de desinformar, sugere que existe uma semelhança entre os processos de informação e processos de desinformação, e que os meios que são normalmente utilizados para informar estão sendo utilizados, também, para desinformar, tornando cada vez mais imperceptível essa “falha no processo de informar” apontada por Floridi (1996, p. 513, tradução nossa).

A outra forma de observar a desinformação como processo é trazida por Skyrms (2010) quando este se refere à **execução de ações sistematizadas (processo) para se atingir o objetivo de desinformar**. O autor aponta que o engano (desinformação)

pode ocorrer a partir de ações sistemáticas (processo) que transmitam informação errônea (coisa) e que beneficiem o emissor em detrimento ao receptor. O próprio Floridi (2012) sugere que a desinformação poderá ser utilizada tanto para se referir às “informações semânticas falsas” quanto ao ato de “induzir em erro seu receptor” (processo). Essa visão trazida por Skyrms e Floridi sugere que, em algumas situações, tão importante quanto a informação de baixa qualidade (coisa) é a forma com a qual os usuários são alcançados por essa informação e como isso pode ser utilizado por eles (processo).

Tudjman e Mikelic (2003) questionam se a desinformação realmente existe como um assunto da prática de documentação da informação. Segundo os autores, a CI não pode negligenciar o fato de que a troca de informações é uma relação social cujos participantes muitas vezes têm interesses opostos. Estes atores se esforçam para mudar as intenções, os objetivos, decisões e ações – não apenas dos envolvidos diretos, mas também de outros participantes indiretos com os quais possam estar em algum tipo de conexão. A visão proposta pelos autores sugere que uma divergência nos interesses que resultam nas ações de desinformação, ou seja, as motivações (desinformação como propósito), levam a utilização de métodos sistemáticos para se atingir os resultados desejados (desinformação como processo).

Skyrms (2010) também aponta o caráter natural de desinformar. Segundo o autor é bastante comum encontrar na natureza seres vivos que se utilizam do processo de desinformação para conseguir a sua sobrevivência. Camuflagens, mudanças anatômicas, emissão de sons e odores, elaboração de rastros falsos, dissimulação comportamental, entre outros, são estratégias utilizadas pelos seres vivos para que possam continuar a existir em seu ambiente. Aqui o autor sugere que os processos de desinformação não precisam ser complexos ou tão engenhosos, e que não são provenientes, necessariamente, de uma visão maligna e destrutiva de um ente, mas que pode ser, inclusive, um processo de proteção. Ou seja, que a desinformação pode ser necessária para equilibrar uma relação de desigualdade ou desvantagem.

Unindo as ideias de Tudjman e Mikelic (2003) e Skyrms (2010), pode-se compreender que a desinformação é um fenômeno natural e social, que não é, necessariamente, desempenhada por seres com extrema capacidade intelectual e que

podem se utilizar de mecanismos simples para proporcionar o engano. Esta noção de que entidades menos inteligentes possam adotar processos simples para atingir seus objetivos quebra um paradigma de que a desinformação só ocorra mediante às estruturas e estratégias complexas de engodo.

Carvalho (2001) afirma que a maior parte das pessoas instruídas não conhece exatamente o que é desinformação. Estas imaginam que este fenômeno se trate apenas de informações falsas para fins gerais de propaganda, ignorando por completo que o principal foco são ações perfeitamente calculadas para uma finalidade específica.

Demo (2000) apresenta um dos processos de desinformação que aconteceu na rede, ao afirmar que a explosão da Internet no início deste milênio expunha um fenômeno trajado da apresentação de uma grande massa de informação, fenômeno mais tarde chamado de *big data*, mas se esquecia da impossibilidade de controle desses dados resultando no sentido oposto da desinformação ou “informação efêmera”.

Ainda segundo o autor, este processo consiste na disponibilização de uma considerável massa de informações de forma contínua e sistemática. Entretanto, o conteúdo oferecido é de baixo valor, impossibilitando ao indivíduo que este passe a conhecer sua realidade e o torne inapto a participar do processo político e tomar decisões relativas ao progresso de sua própria vida e de seus semelhantes. Assim, a enxurrada de informação trazida pela Internet se tornou o método mais refinado no processo de “imbecilização” e que “desinformar será, portanto, parte fundamental do processo de informação” (DEMO, 2000, p. 39)

Nelson e Taneja (2018) apontam que as principais mídias sociais que tentam coibir a disseminação de conteúdo falso, hoje já estão detectando estes problemas não somente pela análise do conteúdo, mas pelo processo como uma mensagem chega a um usuário. O *Whatsapp*, por exemplo, já identifica um conteúdo suspeito somente pela forma como ele trafega entre os grupos, mesmo que o sistema não tenha certeza sobre o assunto que se refere a mensagem. Karlova e Fisher (2013) elaboraram um modelo de disseminação da desinformação a partir do comportamento informacional dos usuários.

Os últimos aspectos a serem considerados nesta visão de desinformação como processo é a ausência de informação e a censura. Iniciando por este último, Floridi (2015)

aponta que a censura pode fazer parte de uma estratégia intencionada a desinformar. A clara negação, a partir de uma estrutura de poder, a outrem para o acesso à informação é um procedimento comum para evitar que esta possa trazer algum tipo de desvantagem ao interessado.

Já tratando da ausência, parece claro que em caso de inexistência da informação, os processos infocomunicacionais que a envolvem ficam comprometidos. Entretanto, não são as óbvias disfunções decorrentes da falta de informação que serão tratadas, mas, o que decorre a seguir, o comportamento humano e suas ações (processos informacionais) a partir desta lacuna.

Hilbert (2012) aponta que em situações em que pessoas acessam informações imperfeitas ou ausentes, o cérebro humano preenche essas “lacunas informacionais” com um conteúdo próprio (depende de cada pessoa) e que tenha significado para o indivíduo. O intuito disso é fazer com que aqueles vazios não venham a se converter problemas em um ponto futuro. O autor usa a teoria do viés cognitivo para explicar este fenômeno. O viés cognitivo é um padrão de distorção de julgamento que ocorre em situações particulares, levando à distorção perceptual, julgamento pouco acurado, interpretação ilógica, ou o que é amplamente chamado de irracionalidade (HASELTON; NETTLE; ANDREWS, 2005).

Hilbert (2012) afirma que existe uma limitação no nosso cérebro em processar informação, e sempre que essa capacidade é extrapolada o cérebro “usa um atalho” para processar esse fato com o intuito de diminuir o esforço cognitivo. É na utilização desses atalhos que a informação é, em parte ou no todo, perdida e se gera os vazios na mente humana. Essas lacunas fazem com que o indivíduo tenha a sua capacidade de tomar decisão prejudicada, e que eventualmente, venha a realizar ações incoerentes com aquilo que é esperado.

Assim, é apresentado mais uma forma da desinformação como processo que são as tomadas de decisão humanas enviesadas decorrentes da ausência ou imperfeição da informação obtida. Ressalta-se que, aqui não se destaca a inexistência/falha da informação em si, mas, o processo infocomunicacional que vem a seguir e como esse pode ser prejudicado por este tipo de limitação biológica.

Sumarizando, a desinformação como processo ocorre em pelo menos duas formas: (i) **desinformação como resultado de um defeito no processo de informar** e (ii) **execução de ações sistematizadas (processo) para se atingir o objetivo de desinformar**. Assim como a desinformação como coisa, essas categorias não são excludentes entre si, e que, em casos específicos, pode-se observar a ocorrência simultânea de mais de uma dessas formas. Observa-se que a desinformação como coisa e como processo normalmente são complementares. O primeiro trata daquilo que será repassado (objeto) enquanto o segundo trata da maneira (mais eficaz possível) de repassar aquela informação para atingir um objetivo (informação como propósito).

2.1.3 Desinformação como Propósito

Por fim, faz-se necessário observar como a desinformação é empregada para atingir algum tipo de finalidade ou atender algum tipo de intenção aqui posta como **Desinformação como Propósito**. Antes de tudo, é necessário apontar que nem sempre a desinformação tem objetivos destrutivos ou maliciosos e que, também, nem sempre os resultados do processo de desinformar são perniciosos.

Para tanto, vamos retomar a sugestão de Tadjman e Mikelic (2003) de que qualquer troca de informação é uma relação social na qual os atores podem apresentar interesses distintos. Então, naturalmente os envolvidos agem para obter um maior benefício da informação (ou desinformação como coisa) ou do momento/forma em que esta é usada (desinformação como processo). Segundo os autores, isso ocorre nos mais variados relacionamentos, desde aqueles do dia a dia dos indivíduos até nas negociações que envolvem membros de altos escalões nas organizações.

O interessante desta visão é que a desinformação é desassociada, em certo grau, das relações de poder onde exclusivamente grandes corporações/organizações se utilizam da desinformação para determinados fins, mas esta também pode ser observada em situações cotidianas, como a conversa de muro entre vizinhos, ou o bate-papo caloroso nos botequins.

Outro ponto que evidencia que a desinformação é algo natural vem da abordagem de Skyrms (2010), na qual este se utiliza do termo engano (*deception*) para descrever a

desinformação como uma questão biológica em que plantas e animais se utilizam de artifícios (tais como camuflagem) no intuito de enganar suas ameaças/vítimas para aumentar suas possibilidades de sobrevivência. Para o autor, o camaleão que muda de cor se utiliza da desinformação para sua defesa, e conseqüentemente, sua sobrevivência. Nesta visão, o autor sugere que não é necessária uma inteligência superior (humana), ou mesmo, a existência de uma estratégia complexa destinada ao engodo para se configurar a desinformação.

Devido à essa ampla possibilidade do que pode ser desinformação, Fallis (2015) explicitou que existem 14 tipos de desinformação, listadas no Quadro 1, a seguir:

Quadro 1 - Tipos de Desinformação por Skyrms

Tipo	Intencional	Leva ao Engodo?
Mentira Maliciosa (ML)	Sim	Sim
Desinformação Visual (VD)	Sim	Sim
Desinformação da Verdade (TD)	Sim	Sim
Desinformação Colateral (SD)	Sim/Não	Sim
Desinformação Adaptativa (AD)	Sim	Sim
Desinformação Altruística (AL)	Sim	Sim
Desinformação Prejudicial (DE)	Sim	Sim
Declarações Verdadeiras (TS)	Sim	Não
Falsidade Acidental (AF)	Não	Sim
Piadas (J)	Sim	Não
Comentários Sarcásticos (SC)	Sim	Não
Verdades Acidentais (AT)	Não	Não
Mentiras Implausíveis (IL)	Sim/Não	Não
Sátira (S)	Não	Sim

Fonte: Adaptado de Fallis (2015)

Para o autor, a **Mentira Maliciosa** é aquela onde há intenção de desinformar, e o principal intuito é diminuir a credibilidade ou reputação da vítima de modo a destruí-la ou diminuí-la moral, emocional ou materialmente. Este tipo de mentira, normalmente, resulta em desinformação e a intenção primária do emissor é prejudicar um alvo (ou grupo de alvos) específicos.

Já a **Desinformação Visual** é aquela que leva ao erro a partir de um dos sentidos, a visão. A forma mais comum desse tipo de desinformação é a camuflagem, entretanto, não está limitado somente a esta, podendo se referir a qualquer engano de natureza

visual. Segundo o autor, esta forma sempre se dá em carácter intencional e sempre resulta em desinformação.

A **Desinformação da Verdade** é aquela onde o emissor se utiliza de uma verdade, em um contexto diferente do original, para fazer com que o receptor mude de opinião, intenção ou mesmo de atitudes. Para o autor, este tipo de desinformação é intencional e resulta em algum tipo de tomada de decisão equivocada por parte de quem recebe a desinformação. Entretanto, Floridi (1996) afirma que este tipo de desinformação se apresenta muitas vezes de forma involuntária (não intencional), principalmente ao considerar que o ambiente da Internet permite a criação, manipulação, reprodução e disseminação de conteúdo de maneira mais ágil, instantânea e, em casos de falsidade, de forma involuntária.

Fallis (2015) também apresenta a **Desinformação Colateral** como aquela que resulta no alcance de um grupo maior do que os receptores originais da mensagem falsa. Este tipo de desinformação pode ser tanto intencional, ou seja, o remetente previa que o conteúdo enganoso iria se propagar além dos destinatários, como também pode ser não intencional e inclusive indesejado pelo emissor. Este tipo de desinformação sempre é um resultado de outra desinformação de algum outro tipo. Ou seja, uma mentira maliciosa pode ter sido usada como a desinformação original, e esta mentira foi disseminada em mídias sociais criando um efeito colateral da desinformação original.

A **Desinformação Adaptativa** (ou evolucionária) é aquela que evolui, se altera e se modifica de modo a sempre beneficiar, de forma sistemática, o emissor. Aqui é possível se referir às mentiras que se perduram e ao longo do tempo vão se transformando para manter o (s) emissor (es) envolvido (s) em uma trama cujo intuito é prover algum tipo de vantagem a eles próprios. Segundo o autor este tipo de desinformação é sempre intencional e tem como resultado o engodo aos receptores da informação.

A **Desinformação Altruística** é aquela que é realizada em benefício do emissor. Este tipo de desinformação é muito comum em casos de saúde, onde expor a verdade ao paciente, ou a todos os familiares, pode causar mais transtornos do que benefícios. Segundo o autor, estas ações são sempre intencionais, resultam no desconhecimento, mas, claramente, não tem objetivos destrutivos ou maliciosos.

Já a **Desinformação Prejudicial** é aquela em que há uma intenção de incentivar algo prejudicial. Ela se difere da mentira maliciosa porque a intenção não é prejudicar um alvo, mas sim trazer ações ou comportamentos prejudiciais à tona. Então, os boatos a respeito das vacinas (incentivando pessoas a não as tomá-las) e das questões climáticas (negando o aquecimento global) são exemplos desta desinformação. Para o autor, esta desinformação se caracteriza pela intenção de prejudicar uma comunidade ou grupo, e existe alguém engajado nesta tarefa. Segundo o autor esta é uma ação intencional e que resulta em desinformação.

Os outros sete tipos de desinformação que vêm a seguir são classificados pelo autor como manifestações “claramente de não desinformação”, mas que podem resultar em enganamentos ainda assim.

Desta forma, o autor considera as **Declarações Verdadeiras** como uma forma de desinformação. Embora na normalidade dos casos a verdade não resulte em enganamentos ou vantagens, existem alguns episódios em que a verdade pode ser mal interpretada pelo receptor. Rogers *et al.* (2016) afirma que essa situação se chama “*Paltering*”, a qual se apresenta como uma forma ativa de fraude que envolve o uso de declarações verdadeiras para transmitir uma impressão equivocada.

Fallis (2015) afirma que esta é uma situação na qual o emissor tem a intenção de fazer tal declaração, mas que não resultaria em desinformação. Entretanto, Rogers *et al.* discordam desta posição, pois segundo eles:

Embora a motivação subjacente para enganar um alvo possa ser a mesma, o *paltering* é diferente de mentir por encomenda e mentir por omissão. Ao contrário da mentira por omissão, o *paltering* envolve o uso ativo de afirmações e, ao contrário da mentira por encomenda, a permeabilidade envolve o uso de afirmações verdadeiras. Assim como a mentira por omissão, o *paltering* pode envolver a falta de divulgação de informações relevantes, mas ao contrário da mentira por omissão, o *paltering* envolve a divulgação ativa de informações verdadeiras, mas equivocadas: a manipulação permite que os possíveis enganadores influenciem ativamente as crenças de um alvo (ROGERS *et al.*, 2016, p. 461).

Assim, podemos compreender que existem pontos de vista que defendem que as declarações verdadeiras podem resultar em alguma espécie de desinformação.

Já a **Falsidade Acidental** se dá quando o emissor não está intencionado em mentir, mas age no sentido de desinformar. Segundo Fallis (2015), este tipo de ação não

recai no comportamento de desinformação intencional (*disinformation*), mas, certamente recai na classificação de não intencional (*misinformation*). Os casos mais comuns de falsidade accidental são as mensagens compartilhadas nas mídias sociais sem a checagem por parte do usuário.

Piadas e Comentários sarcásticos entram nesta categoria, uma vez que estes normalmente tratam de conteúdo inverossímil, situações inverídicas ou de declarações com sentido inverso ao real. Normalmente as piadas e os comentários sarcásticos não têm a intenção de enganar e são feitos de forma premeditada pelo emissor.

Já a **Verdade Accidental**, na visão apresentada pelo autor, se dá quando o emissor está intencionado em dizer algo falso, que por circunstâncias diversas é uma verdade cuja revelação não é a intenção de quem emite. O exemplo dado por Fallis (2015), é uma situação em que alguém ao ser questionado pela polícia do paradeiro de um amigo, a pessoa, acreditando que o amigo está fora da cidade com os primos, afirma que o amigo está escondido no cemitério com a intenção de enganar a polícia. Entretanto, sem o conhecimento do interrogado, o amigo realmente está no cemitério. Ou seja, Fallis não considera aqui as verdades, desconhecidas até então, reveladas de forma accidental, mas aquelas cujo emissor elabora uma mentira para enganar o receptor, mas que acidentalmente este engodo acaba por se revelar como verídico.

Floridi (2011) questiona se a verdade accidental, nos moldes apontados por Fallis (2015), pode ser considerada desinformação, uma vez que a falsidade não se concretiza. Entretanto, quando a verdade accidental resulta na revelação de fatos desconhecidos, normalmente se dá na quebra da desinformação altruística. Floridi (2011) sugere que a falha no processo de informar se configura na desinformação, já que a verdade foi revelada de um modo não desejado.

As **Mentiras Implausíveis** são aquelas que se configuram como absurdas a ponto de não possuir credibilidade suficiente para enganar o receptor. Este tipo de desinformação é comum nas histórias contadas a crianças (ex. Papai Noel, Fada do Dente) onde estas acreditam em situações fantásticas que são facilmente interpretadas como ficção por adultos. A plausibilidade da mentira ocorre pelo bom senso do emissor (daí a questão de que esta pode ou não ser intencional) quanto pelo discernimento do receptor. Normalmente, tais situações não resultam no engano do receptor.

Por fim, as **Sátiras** também se referem às situações inverídicas criadas para sugerir situações semelhantes em contextos que estão relacionados. Normalmente a intenção da sátira não é o engodo e são feitas de forma intencional pelo emissor.

Observando todas estas formas de desinformação, é possível inferir que mais de uma delas podem acontecer no mesmo momento. Por exemplo, em casos de golpes uma mentira (maliciosa) pode ser espalhada para um grande número de pessoas (colateral) para benefício sistemático de um terceiro (adaptativa). É possível observar também o que Fallis (2015) ressalta a respeito da intencionalidade do emissor e se o desfecho resulta em erro, engano ou engodo.

No primeiro ponto, Fallis (2015) tratou de estabelecer a desinformação como informação não acidentalmente enganosa e a este fenômeno ele chamou de *disinformation*. Entretanto, muitos autores também compreendem que nem sempre a desinformação ocorre de maneira voluntária (erro honesto). E essa intencionalidade será o ponto de partida da construção do raciocínio da desinformação como propósito, uma vez que, por se tratar de um fenômeno que também é social e natural, aparentemente, só é possível compreender um objetivo para a desinformação caso esta ocorra de maneira premeditada, planejada ou direcionada a um fim.

Fallis (2015, p. 401) trata a desinformação (*disinformation*) como “informação enganosa que tem a função de enganar”. Para esse autor, o erro honesto, a negligência e o preconceito inconsciente não se confundem com a desinformação que representaria engano intencional e informações imprecisas, com o propósito de enganar as pessoas. O autor atribui a esta desinformação o nome “*misinformation*” para separar esta daquela que ocorre de forma intencional (*disinformation*). Fallis (2015, p. 401) reforça essa ideia ao afirmar que a “desinformação é uma informação imprecisa que a fonte sabe ser imprecisa”. Nos próximos parágrafos, será utilizado o termo *misinformation* (em inglês) para diferenciar esta da desinformação proposital e evitar problemas em questões de tradução.

Também reconhecendo a *misinformation*, Fetzer (2004) afirma que a desinformação ocorre apenas nas ações que deliberadamente resulta no engano de outrem, ou seja, o erro seria uma ação involuntária que não se confundiria com o tipo de desinformação proposital. Já Floridi (2011, p. 260) afirma que a “desinformação é

simplesmente uma má informação propositadamente transmitida para enganar o receptor, fazendo-o acreditar que é informação”. Essa visão também é corroborada por Pinheiro e Brito (2014), bem como Zattar (2017) quando estes afirmam que desinformação é engano proposital.

É possível perceber que há um reforço sobre a presença da intencionalidade para existir a desinformação. Entretanto, muitos autores não afirmam que em casos nos quais não há a intencionalidade, não se trate de casos de desinformação, embora haja o erro.

Fallis (2015) aponta claramente como *misinformation* o erro honesto, a negligência e até mesmo o viés cognitivo. Floridi (2005) afirma que “quando o conteúdo semântico é falso, este é um caso de *misinformation* [...] E se a fonte da desinformação é consciente de sua natureza, pode-se falar de desinformação”. O mesmo Floridi (2011, p. 251, tradução nossa) afirma que “desinformação é simplesmente a *misinformation* intencionalmente transmitida para enganar o receptor em acreditar que é informação”. E Skyrms (2010) aponta a intenção, não só cognitiva, mas biológica (evolução e adaptação das espécies) para desinformar.

Essa referência à “não intencionalidade” é incomum na Ciência da Informação onde há reforço positivo, ou seja, que a desinformação provém de um propósito, é o mais comum. Os trabalhos de Fallis (2015) e Floridi (1996; 2011) se mostram importantes neste sentido por evidenciar que a *misinformation* também podem resultar em erros, enganos, engodos, decisões e julgamentos errôneos e tomadas de decisões equivocadas, merecendo assim ser investigada tanto quanto a desinformação.

O próprio termo *misinformation* sofreu uma transformação dentro da própria Ciência da Informação. Os primeiros trabalhos publicados da área que adotaram este termo são da década de 1950, e naquele momento se referia aos defeitos no processo de informar, ou seja, quando a tentativa de informar não era realizada com êxito. Esta forma de enxergar a *misinformation* ainda vigorava na década 1990, uma vez que o próprio Capurro (2003) ainda usa o termo para falar dos valores negativo da informação.

O primeiro trabalho que apresenta de forma separada os dois termos foi dado por Dretske (1983), mas somente com os trabalhos de Hernon (1995) e Floridi (1996) que os termos *misinformation* e *disinformation* passaram a ter a conotação atual. Vale ressaltar a compreensão de que o termo *misinformation* evoluiu, deixando de existir em seu

sentido original ao se referir às falhas no processo de informação (desinformação como processo) para se tornar algo relacionado à desinformação não intencional (desinformação como propósito).

Por fim, destaca-se que embora o termo *misinformation* tenha passado por esta transformação, ele ainda é tratado em pesquisas contemporâneas com o seu sentido original de falhas no processo de informação. Isso significa que muitos dos trabalhos em língua inglesa que se referem à desinformação como processo ainda utilizam esta nomenclatura para se referenciar a desinformação. E é por isso que a pesquisa que envolve este termo pode se tornar confusa, uma vez que são observados aspectos diferentes da desinformação.

Entretanto, não é sobre a *misinformation*, em seu sentido hodierno, que se trata a desinformação como propósito, e sim sobre a desinformação não acidental que possui um objetivo a ser alcançado pelo emissor. Não são abundantes as referências na Ciência da Informação sobre as intenções em desinformar, mas, dentre o pouco encontrado, existe uma discussão sobre três situações: **(i) Diminuir a capacidade de reflexão dos indivíduos; (ii) Implementação de projetos de poder; e (iii) Alertar para a impossibilidade de acesso à informação de qualidade por parte das maioria das pessoas.** Ressalta-se que, em grande parte dos trabalhos, estas situações não são excludentes, muito pelo contrário, normalmente ocorrem de forma conjunta para se atingir os objetivos.

Quanto a **diminuição da capacidade de reflexão dos indivíduos**, é possível retomar a perspectiva de Loftus (2015) quando este trata da “transformação mental” do indivíduo e da assimilação do conteúdo inverídico como verdade. Desta forma, a autora sugere que a desinformação mais do que criar falsas memórias, torna estas inverdades o estado real para uma pessoa. Fallis (2015) também corrobora esta visão ao afirmar que desinformação é a informação capaz de construir “crenças falsas” levando o indivíduo a refletir e agir sob a influência de falsas perspectivas.

Além das questões de vieses cognitivos outros autores citam o processo da diminuição da capacidade dos indivíduos em discernir a realidade ou tomar decisões, uma vez que estes possuem acesso à informação de baixo valor cultural ou formativo. Alguns autores dão nomenclaturas diferentes para este processo, por exemplo, Demo

(2000) e Pinheiro e Brito (2014) o denotam por “imbecilização”, já Leite e Matos (2017) chamam de “zumbificação”, Alves (2001) nomeia de “emburrecimento”, e por fim, Flinders e Wood (2015) chamam de “despolitização”.

Pinheiro e Brito (2014) aponta que indivíduos menos esclarecidos se encontram em um estado de “precariedade informacional”, resultando em uma “ausência de cultura” e incompetência informacional que dificilmente será revertida, uma vez que o próprio interagente não consegue localizar a informação (de valor agregado) que deseja. Os autores também ressaltam que essa falta de competência informacional resulta na manipulação, em grande parte, de material de baixo nível cultural, mantendo esse ciclo de “imbecilização” em determinados setores sociais.

Demo (2000) promove um rico debate acerca deste processo chamado por ele de “cultivo da ignorância”. Para o autor, isso não se dá pela falta de informação, mas justamente pelo oposto, a partir do excesso de dados (onde a maioria é de baixa qualidade) em um contexto onde se torna difícil para o usuário se nutrir de informação útil e de valor. O mesmo autor sugere o surgimento da “sociedade da desinformação” construída de forma acelerada pelas redes (Internet), que é formada por uma maioria de pessoas que se enxergam como cultas, mas que as informações disponibilizadas cotidianamente não oferecem nenhum tipo de conhecimento para transformar, ou mesmo alicerçar sociedades.

O autor afirma que a informação habilita o indivíduo ao pensamento crítico, questionador, capaz de construir e desconstruir realidades. Enquanto que o processo de desinformação pode tornar o sujeito incapaz, incompetente ou mesmo inábil em reproduzir este comportamento de reflexão. E ainda, quando esta habilitação se dá para propósitos suspeitos, obscuros e que está a serviço de um grupo pequeno de pessoas em forma de doutrinação. Neste contexto, o conhecimento pode ser deturpado para elaborar o que o autor chama de “processo de imbecilização”.

Quanto ao segundo ponto que se trata do **uso da desinformação como estratégia para a implementação de projetos de poder**, pode-se retomar o próprio contexto originário da palavra (дезинформация) onde governos se utilizavam deste artifício para conseguir uma melhor posição política.

Nunes *et al.* (2018) mostraram como a desinformação foi utilizada (automatizada) para manipular processos eleitorais. Os autores perceberam o caráter global das ações, e que grupos de pensamento ideológico semelhante se beneficiaram da mesma *botnet* para a campanha de desinformação. Assim, concluíram que não há como se analisar o impacto eleitoral deste processo, e conseqüentemente, medir a efetividades das ações, entretanto, sugerem que o repetido uso da estratégia (duas vezes no Brasil e duas vezes EUA) a torna interessante e deve trazer algum tipo de retorno.

Carvalho (2001) aponta que a parcela da sociedade que possui competência informacional e que também tem acesso à informação de qualidade acredita que a desinformação só ocorre em meios restritos com finalidades comerciais (propaganda), não considerando que este processo pode estar atrelado a interesses maiores, com objetivos específicos para que o comportamento desta sociedade seja mais favorável aos interesses dos manipuladores da opinião.

Demo (2000) novamente se destaca nestas reflexões ao afirmar que a desinformação pode ser o meio pelo qual “pequenos grupos” realizam seus “projetos colonizadores”. Entretanto, ao contrário da estratégia normalmente concebida onde a falta de informação e censura é a maior forma de desinformação, o que se vê é que a própria forma de se observar a informação e seus modelos (como a da alcunha sociedade da informação) são utilizados de forma antagônica aos seus propósitos. Assim, a informação benéfica estaria mesclada com aquela cujo propósito é desinformar. Segundo o autor, a informação seria a base de um “refinado processo de imbecilização” e que desinformar será “parte fundamental do processo de informação” (DEMO, 2000, p. 39).

Demo aponta que os atores por trás desta grande máquina da desinformação são complexos e de natureza variada. Os articuladores de tais projetos se utilizam dos meios de comunicação tradicionais e das redes digitais para massificar essa informação de baixo valor.

Por fim, Pinheiro e Brito (2014) sugerem que existe algum tipo “elite” que se utiliza também da desinformação como estratégia para perpetuação de poder. Ao desinformar a sociedade, essa elite teria uma maior facilidade de atingir seus interesses.

Por último, destacam-se as menções a impossibilidade de acesso à informação de qualidade por parte da sociedade. Neste quesito, existe uma preocupação dos autores na existência de um conjunto de pessoas que não têm acesso às informações básicas que podem tornar a vida delas mais digna e humana. Pinheiro e Brito (2014) apontam que existe uma divisão entre informados e desinformados, na qual os últimos representam aqueles “excluídos do acesso a bens informacionais”.

Aquino (2007) aponta as consequências desta desinformação na vida dos indivíduos ao se referir que estes, ao se relacionar com outras pessoas e o mundo, estão em um “estado de desconhecimento” total ou parcial que se configura em todas as ações da vida destas pessoas.

Questões de letramento e competências informacionais são discutidas com riqueza na Ciência da Informação, mas pouco relacionadas ao tema desinformação. Tais situações são mais discutidas em um conceito social, onde a vida do indivíduo se torna prejudicada pela impossibilidade do bem informacional, por mais básico que este seja. Pode-se afirmar não haver muita discordância de que a deficiência de letramento e competências informacionais resulte invariavelmente, em desinformação, mas o diálogo sobre esta ótica ainda precisa de um maior aprofundamento.

Surpreendentemente, a desinformação como propósito é aquela em que, aparentemente, existe um maior espaço para a discussão. Aspectos sociais, políticos, comerciais e informacionais que envolvem um sem fim de temas (competência e letramento informacional, por exemplo) e de atores são vislumbrados. Mas, que grande parte do debate reside na intencionalidade (*misinformation x disinformation*) do emissor da informação.

Compreende-se, assim que esta é, das três facetas da desinformação, aquela que mais apresenta lacunas e perguntas a serem respondidas e ações a serem propostas.

Observar a desinformação a partir de suas facetas (coisa, processo e propósito) auxilia a compreensão de que esta é um fenômeno amplo e que se aplica a diversas situações cotidianas, capaz de modificar relações e estruturas sociais das mais capilares as mais poderosas. Amplo, porque é um fenômeno natural (animais que são especialistas em desinformar), social (pode acontecer no bate papo da esquina), que

pode se dar a partir de informação verdadeira (desinformação da verdade) e para auxiliar outrem (desinformação altruística).

Um objeto de estudo desta dimensão precisa de um olhar, embrionário neste momento, sobre seus aspectos e suas características nos mais diversos ângulos e pontos de vista. Reforço, que a sugestão destas três facetas, não as torna definitivas, muito pelo contrário, são um ponto de partida na criação de um modelo em que novos olhares sobre a desinformação encontrem um alicerce para prosperar.

Se as facetas envolvem a forma de se olhar para a desinformação, como é que estas ocorrem no cotidiano das pessoas? Será que ela se manifesta de uma só maneira ou, assim como as facetas, ela também possui mais de um meio de ocorrer? No intuito de responder a estas perguntas é que foi desenvolvida a próxima subseção de manifestações da desinformação.

2.2 Manifestações da Desinformação

Schrader (1986, p. 180) materializa a desinformação em "mentiras, propaganda, deturpação, rumores, alucinação, ilusão, erro, ocultação, distorção, embelezamento, insinuações e enganos". Apesar de clara a visão do autor sobre quais ações estão ligadas à desinformação, esta apresenta dois problemas: (i) como comparar questões que não se referem às mesmas classes situações (ex: mentiras se refere à informação inverídica, enquanto que ocultação se refere ao ato de ocultar, então como comparar os dois?); e (ii) como estes abarcam as manifestações mais recentes de desinformação tais como notícias falsas e pós-verdades.

Assim, faz-se necessário observar como este fenômeno ocorre na vida real, principalmente na Internet. Desta forma, foram identificadas sete manifestações de desinformação que acontecem na rede.

2.2.1 Pós-Verdade

Iniciando pela pós-verdade, esta é uma das manifestações mais presentes no contexto da desinformação. Côrrea e Custódio (2018, p. 2) afirmam que dentro da atual

cultura digital este fenômeno se apresenta com mais força, tendo em vista que em ambientes como a Internet é comum a disseminação de inverdades, em virtude do compartilhamento de conteúdo de forma irrestrita e indiscriminada. Destacam os autores que no momento atual se vivencia a era da pós-verdade, que se justifica pelo “compartilhamento ininterrupto e indiscriminado de informações que transforma a Internet em um ambiente onde ‘inverdades’ se espalham com frequência e mais rapidamente do que os fatos reais” (CÔRREA e CUSTÓDIO, 2018, p. 2).

Para o Dicionário de Oxford¹, pós-verdade (*post-truth*) teria o seguinte significado: “Relaciona-se ou denota circunstâncias nas quais os fatos objetivos são menos influentes na formação da opinião pública do que os apelos à emoção e à crença pessoal”. Apesar dos avanços da academia no desenvolvimento de pesquisas sobre o tema, estes ainda estão focados em explicar situações que envolvem a vinculação de notícias falsas nas redes sociais virtuais, tanto que em trabalhos nacionais da CI não foram encontrados, em termos de significado, do que viria a ser pós-verdade.

Silva Filho, Silva e Luce (2017, p. 271) destacam que o termo pós-verdade refere-se a um fenômeno recente, escolhido pelo Dicionário Oxford, em 2016, como a palavra inglesa que mais se destacou durante o ano, apresentando um aumento em seu uso de 2.000% no período. Os autores sugerem que este aumento de notoriedade se deu devido ao aumento de problemas relacionados à “disseminação de informações falsas, não-cheçadas, boatos, calúnias, difamações, entre outros, nas mídias sociais, principalmente no Facebook, Twitter e WhatsApp, que possuem um vasto alcance e influência no usuário comum da Internet”. Segundo os autores, este mesmo fenômeno de pós-verdades é um dos catalizadores da polarização vista na Internet contemporânea.

Bezerra, Capurro e Schneider (2017) afirmam que a humanidade vive em “um estado de pós-verdade ou semiverdades, clichês, mentiras e desinformação”. Os autores atribuem a esta era uma realidade similar àquela em que Heidegger afirma que o “Homem, ou pensar (e ser) o que todos pensam (e são) sem pensarem por si mesmos, sem se atreverem a ser eles mesmos, aderir a esses todos pensam (e são) assim”. A percepção do efeito manada, antigamente liderados pelos meios de comunicação de massa, é agora observada também nas redes sociais.

¹ <https://en.oxforddictionaries.com/definition/post-truth> <acesso em: 30/08/2018>

A preocupação dos autores é que a realidade atual, quando se fala de pós-verdade, chegue ao ponto de uma luta de poderes midiática, política e econômica nos moldes de um “regime de verdade”, de Foucault. Esses regimes de verdade são “entendidos como conjuntos ordenados de proposições, instituições e disciplinas que organizam e controlam os discursos e impõem-se como estratégias de manutenção do poder, por meio de uma política universal da verdade submetida às disciplinas e sanções” (BEZERRA; CAPURRO; SCHNEIDER, 2017, p. 374).

Nesse sentido, é possível compreender que a pós-verdade pode se apresentar enquanto desinformação como **processo**, como **coisa** e como **propósito** de forma conjunta, já que o emissor pretere o fato em si (informação como coisa) e dá prosseguimento a opiniões ou crenças. Podem existir casos em que a opinião/crença do indivíduo represente a informação em si, mas, normalmente, ou informações de baixa qualidade ou fora de contexto (processo) são disseminadas.

Há pós-verdade mesmo quando se existe um contexto onde o emissor não tem a intenção de desinformar, este se encontra em certo estágio de ignorância e repassa informação baseada em suas crenças, o que recai na transformação mental (imbecilização ou emburrecimento) do receptor.

2.2.2 Notícias falsas

Outro fenômeno que está intimamente ligado à desinformação são as notícias falsas (*fake news*), termo comumente utilizado para se referir aos acontecimentos que desencadeiam estados de desinformação, e que muitas vezes são empregados de forma equivocada como sinônimo desta razão, pela qual merece destaque o fato de que nem toda desinformação é promovida através da disseminação de uma notícia falsa, podendo aquela apresentar-se por outras formas.

Segundo Zattar (2017, p. 286), as notícias falsas devem ser enxergadas como uma prática informacional onde reside “questões que envolvem a qualidade do conteúdo nas dinâmicas de busca e recuperação, dentre as quais estão as notícias e informações falsas ou semifalsas, a desinformação”.

Leite e Matos (2017, p. 2336) relacionam a produção de notícias falsas com o aumento do uso da Internet, destacando que:

Esse tráfego de informações por meio de novas formas de acesso e produção de conteúdo, porém, tem possibilitado o consumo e disseminação de informações falsas, distorcidas, manipuladas, servindo às mais diversas finalidades pessoais e institucionais. A popularização de termos como 'fake news', 'pós-verdade' e 'desinformação' tem trazido à tona uma recente preocupação com a veracidade e a confiabilidade das informações disseminadas na web, as quais acabam formando opiniões e construindo pretensos conhecimentos, baseados em informações falsas ou imprecisas (LEITE; MATOS, 2017, p. 2336).

Os autores exemplificam a presença deste fenômeno na Internet a partir do número de compartilhamentos, nas redes sociais, de notícias falsas sobre as eleições americanas de 2016 e o caso conhecido como 'Operação Lava-Jato' no Brasil. E afirmam: "o número de interações nas redes sociais com as notícias falsas excedeu o de interações com as notícias que, de fato, eram verdadeiras" (LEITE; MATOS, 2017, p. 2336). Esse acontecimento é denominado como 'zumbificação da informação' por ser um processo em que a disseminação e o consumo de informações falsas ou distorcidas provocam uma infecção generalizada da desinformação na Internet.

Seguindo esse raciocínio, Silverman e Singer-Vine (2016, p. 1, tradução nossa) observam que as "manchetes de notícias falsas enganam adultos americanos em 75% do tempo". Ademais, 62% dos americanos recebem notícias falsas nas plataformas de mídias sociais (GOTTFRIED; SHEARER, 2016), corroborando para que sejam mais compartilhadas que as verdadeiras.

Para Shu *et al.* (2017), a questão da disseminação de conteúdo de baixa qualidade, fabricado intencionalmente, ou seja, notícias falsas, ganha maior expressão no ambiente das mídias sociais por estas oferecerem fácil acesso a baixo custo. Sendo assim:

Notícias falsas podem gerar utilidade para alguns consumidores, mas também impõe custos sociais, tornando mais difícil para os consumidores inferir o verdadeiro estado do mundo - por exemplo, tornando mais difícil para os eleitores deduzir quais candidato que eles preferem (ALLCOTT; GENTZKOW, 2017, p. 212, tradução nossa).

Bezerra, Capurro e Schneider (2017) afirmam que robôs são responsáveis por parte da disseminação de notícias falsas. Segundo eles a atual disseminação de notícias

falsas na Internet, é realizada em grande parte por robôs (*bots*), e é realizada em um volume incomensurável de desinformação se tornando um problema para as principais plataformas da Internet usadas para a divulgação de notícias, que criam mecanismos para detectar e bloquear o conteúdo produzido por perfis dedicados à viralização de tais notícias. Burkhardt (2018) afirma que *bots*, *bad bots*, especificamente, já são os maiores responsáveis pela disseminação de notícias falsas. O mesmo entendimento apresenta Chao *et al.* (2017) ao sugerirem que *bots* já assumem um papel de protagonismo na disseminação de *fake news* nas mídias sociais, sendo que a ação desses agentes foi decisiva em processos eleitorais realizados na França e EUA, como será visto no decorrer deste trabalho.

Apesar da proximidade em que se apresentam a pós-verdade e as notícias falsa (*fake news*), é necessário entender que estas não se confundem, pois, apesar de resultar no mesmo processo, fenômeno ou propósito de desinformação, o *modus operandi* delas são distintos. Desta forma, reconhecem Bezerra, Capurro e Schneider (2017) que as notícias falsas se referem ao processo de criação e disseminação de conteúdo inverídico, enquanto a pós-verdade se refere ao comportamento de pôr os fatos em segundo plano e tornar opiniões e pré-conceitos em um patamar de maior relevância.

Sendo assim, é importante destacar que notícias falsas são “informações fabricadas que imitam o conteúdo da mídia na forma, mas não no processo ou na intenção da organização” (LAZER *et al.*, 2018, p. 1094, tradução nossa). Assim, é possível compreender que se trata de um conteúdo que apesar se revestir de certa formalidade quanto à sua aparência, peca no que diz respeito à qualidade das informações que trazem em seu corpo, podendo se dá de maneira deturpada, enganosa, obscura, enviesada.

Aqui é possível ver uma forte correlação desta manifestação com a desinformação como **coisa**, afinal, a notícia falsa, em sua maioria, significa que a informação apresenta imperfeições afetando a sua qualidade. Existem casos onde uma, ou mais, notícias reais são utilizadas em contextos diferentes, e normalmente desfavorável, aos originais e neste caso, se houver uma sistematização do uso dessas notícias falsas, pode-se observar o uso das *fake news* em uma estratégia (desinformação como **processo**) de mal informar.

Também considerar casos onde as notícias falsas se apresentam na estratégia da desinformação como **propósito**. Estas podem resultar na formação cognitiva dos receptores, os quais a partir disto podem tomar decisões ou assumir comportamentos que não sejam adequados a determinadas situações. E também, se pode especular sobre a plausibilidade do uso destas notícias, como uma das “armas”, em uma estratégia de implementação de projetos de poder.

2.2.3 Data Flood (Enxurrada de Dados)

Em 1967 Ackoff apresentava uma crítica às teorias ligadas ao gerenciamento de sistemas de informação ao apontar que estas tentavam trazer ordem ao caos decorrido de estruturas que possuíam falta de informação útil e excesso de informação desnecessária. Ackoff (1967) intitulou sua crítica de “gerenciamento de sistemas de desinformação”, e apesar da discussão sobre a desinformação, àquela época, ser quase inexistente, o autor acabou antevendo um dos principais problemas ligados à desinformação.

Lewandovsky *et al.* (2012) aponta que essa enxurrada de dados se mostra como uma forma eficiente de promover desinformação. No que ele chama de efeito de familiaridade, a partir do momento em que os usuários se tornam cada vez mais expostos ao conteúdo desinformativo, mais familiar e crível esta informação parece ser. Ele aponta para a união deste efeito familiaridade ao efeito manada promovido pelo *Twitter*, como ferramentas poderosas na disseminação de notícias falsas.

Frenda, Nichols e Loftus (2011) apontam para outro efeito causado pela enxurrada de dados que é a disseminação seletiva. Ao lidar com uma grande quantidade de dados, os indivíduos tendem a manter na memória aquilo que lhes é mais interessante ou conveniente, e isso ocorre com muitas das notícias falsas. Então, mesmo quando um fato verídico é apresentado, ele pode se perder no meio do mar de conteúdo que é apresentado ao usuário comum. Corroborando esse fenômeno, Demo (2000) chama a atenção em como o excesso de dados pode ser utilizado como uma estratégia de desinformação.

A expressão *Data Flood*, como também é conhecida a enxurrada de dados, na gênese do seu significado pode ser traduzida como inundação de dados ou compreendida como grande volume de informações. Kietzmann *et al.* (2011) apontaram que “*Data Flood*” ocorre nas mídias sociais quando a criação de conteúdo se dá em quantidade massivas “afogando” os usuários em informações.

Aqui é possível perceber uma forte correlação desta manifestação com a desinformação como **processo**, afinal, a enxurrada de dados se refere à situação em que um indivíduo se expõe a uma saturação de informação, ao ponto que este só consiga observar algumas destas (incompletude) e com uma profundidade menor do que o desejado (superficialidade). Assim, não há uma garantia do correto processo de informar, ou mesmo uma premeditada ação de desinformar.

A desinformação como **coisa** também pode se fazer presente, pois neste cenário existe uma grande plausibilidade de que ocorra na situação o viés cognitivo, entretanto, esta visão só se mostra plausível quando o emissor não tem a intenção de desinformar, e as informações lá constantes fazem parte de um plano para informar. Ou seja, que o processo de mal informar não seja intenção do emissor, ou resultado de uma falha de execução.

Por fim, a desinformação como **propósito** também pode ser observada a partir da intencionalidade do emissor em confundir ou “mal-formar” o receptor. Além desta, sempre se pode apontar o resultado da desinformação que é a transformação mental dos indivíduos que pode levar a má decisões ou comportamentos.

2.2.4 Falsa Propaganda

Frenda, Nichols e Loftus (2011) apontam que a diferença entre falsa propaganda e notícias falsas é muito tênue. A primeira tem o intuito de convencer/converter enquanto a segunda tem como função informar/formar/mal formar. Propagandas não se tratam de conteúdos nem notícias falsas em si, mas são ligadas a distorcer o contexto para levar o espectador a uma conclusão específica, e normalmente equivocada, sobre o assunto.

Lewandovsky *et al.* (2012) sugerem que os elementos mais comuns da falsa propaganda é a omissão/ocultação de informações importantes sobre o contexto ou

apresentar fatos reais que contenham aspectos bons e ruins e só as primeiras são focadas/maximizadas, enquanto que os problemas são esquecidos ou minimizados, e por fim, distorções leves a moderadas de parte do contexto para direcionar o pensamento. Aqui é importante lembrar que o próprio Schrader (1986) evidencia a propaganda como uma das variações da desinformação.

A falsa propaganda possui uma forte correlação tanto com a desinformação como **processo** quanto como **propósito**, uma vez que pode existir uma sistematização do processo de propaganda para se atingir um objetivo específico. Nem sempre a propaganda se manifesta a partir da desinformação como **coisa**, uma vez não se faz necessário o uso de conteúdo inverídico em si.

2.2.5 *Internet Bias* (Viés da Internet)

Frenda, Nichols e Loftus (2011) apontam que esta é a forma mais leve e a que está mais presente na Internet. O *Internet Bias* representa os pequenos desvios comportamentais, ligados a informação em rede, e que, não necessariamente, implicam em um dos quatro outros citados anteriormente. Por exemplo, uma notícia antiga que é postada nas redes sociais, apresentando conteúdo que pode trazer uma outra interpretação na época vigente pode ser enquadrar nesta categoria e não necessariamente ser falso (notícia falsa), ou opinativa (pós-verdade), nem ter sido viralizada (*data flood*) nem representar propaganda falsa.

Lewandovsky *et al.* (2012) sugerem que esses desvios intencionais levam o cérebro a construir a memória sobre o fato de forma mais receptiva (viés cognitivo). Por exemplo, se uma notícia possui verdades e mentiras, o foco na mentira pode torná-la mais plausível. Então estratégias como enfatizar o mito, e não o fato, ou mesmo explicar o mito de forma simples e o fato de maneira mais complexa e aprofundada, mesmo sendo verdadeira, torna o mito mais atrativo.

Fallis (2015) aponta que esta visão de desinformação é muito comum nos usuários que sofrem com comportamentos indevidos na Internet, tais como *Troll* (na linguagem coloquial a trolagem), *Bulling* (humilhação/difamação de um indivíduo), vazamentos (*Leaks*) e *Harassment* (humilhação/difamação de um grupo). Nestes fenômenos ocorre

uma espécie de julgamento sobre o alvo, mas o intuito não é convencer/converter pessoas e sim destruir uma reputação (mentira maliciosa).

Por se tratar de uma gama muito grande de eventos e que envolve polêmicas atuais da rede, tais como o jogo da baleia azul, os vazamentos de dados e os muitos casos de *Bulling* e *Harassment* em rede, existe a visão de que está ligado de uma maneira mais próxima à desinformação como **propósito**.

Nos casos de *Bulling* e *Harassment* é comum também as informações estarem associadas a falsas crenças ou mesmo fatos inverídicos sobre pessoas e grupos de tal modo que não é incomum associar tais práticas aos elementos observados na desinformação como **coisa**.

É mais difícil associar este tipo de manifestação à desinformação como **processo**, uma vez que não são observados, geralmente, processos complexos para se desenvolver tal tipo de desinformação. Entretanto, pode-se discutir a questão do jogo da baleia azul, o qual se trata de uma série de desafios que inicialmente se mostra como coisas inocentes e vai evoluindo a pequenas transgressões, até chegar ao último desafio: o suicídio. Aqui há um processo de ganhar a confiança da vítima e até mesmo tornar o jogo instigante e divertido ao se cruzar linhas morais e éticas (de uma forma leve), mas o objetivo final é levar o jogador ao suicídio. Nesse sentido, pode-se dizer que há, neste caso, um processo bastante elaborado de desinformação.

2.2.6 Social Spam

Esta também é uma das formas de manifestação da desinformação residindo na distribuição de conteúdo indesejado nas redes. Aqui estariam inclusas mensagens com palavras de baixo teor, insultos, discursos de ódio, informações sensíveis sobre usuários, *links* maliciosos e outros. Desta forma, pode-se compreender que o *Social Spam* é a disseminação da informação de forma automática na rede.

Markines, Cattuti e Menczer (2009) apontam que o “Social Spam” é a ação deliberada de espalhar conteúdo, de interesse ou não, em mídias sociais de forma automatizada e com o maior alcance possível. De acordo com os autores, as principais

motivações para o uso do *Social Spam* é o baixo custo, rapidez de disseminação da informação e a considerável capacidade de penetração nas massas.

A ativação do conteúdo *spam* põe o usuário imerso a um volume incomensurável de informações não relevantes para as discussões e debates levantados em determinado contexto provocando a desinformação deste, pois resulta na limitação da sua capacidade de ter uma visão macro e bem informada da realidade, é o que ocorre, por exemplo, na plataforma do *Twitter*, rede em que ocorrem campanhas *spam*, havendo a manipulação de várias contas com o intuito de espalhar informações de maneira automatizada (CHU *et al.*, 2012, p. 455).

Assim como a falsa propaganda, o *social spam* possui uma forte correlação tanto com a desinformação como **processo** quanto como **propósito**, uma vez que pode existir uma sistematização do processo de distribuição das mensagens com o intuito de se converter ou conquistar o maior número de consumidores possíveis para um determinado fim, que nem sempre é um fim lícito (golpes). Nem sempre a mensagem é falsa ou de baixo valor de qualidade, o que enfraquece a relação dela como a desinformação como **coisa**, embora, que nos casos de golpes na Internet a informação inverídica é indispensável para a conclusão bem-sucedida deste processo.

2.2.7 Deepfake

Esta talvez seja a modalidade mais recente de desinformação que surgiu nos últimos tempos, não possuindo uma tradução ou adaptação do termo para a língua português. Destaca-se que “*deepfakes* são simplesmente recentes demais para qualquer um ter muita perspectiva. O software de aprendizado de máquina que os permite surgiu apenas no final de 2017” (FLETCHER, 2018, p. 457).

Desta forma, cabe aqui conceituá-la para melhor compreensão do termo, que deve ser considerado como “técnicas de criação de vídeos falsos hiperrealistas, ameaçam a consistência de informações compartilhadas digitalmente, potencialmente sensíveis” (LARA *et al.*, 2018, p. 1), ou ainda, “vídeos produzidos de pessoas dizendo e fazendo coisas que nunca fizeram” (FLETCHER, 2018, p. 455).

Para Chesney e Citron (2019, p. 1) trata-se de “recentes avanços em uma forma de inteligência artificial conhecida como “*deep learning*”, em que conjuntos de algoritmos chamados “redes neurais” aprendem a inferir regras e replicar padrões examinando grandes conjuntos de dados”, assim, é razoável observar que a *deepfake* poderá ser empregada com fins criminosos ou malignos. Neste contexto, pode-se incluir também a desinformação, posto que se trata de:

Um conjunto de técnicas usadas para sintetizar novos produtos visuais, exemplo, substituindo faces nos originais. Os casos típicos envolvem filmes pornográficos em que os rostos de atrizes famosas como Gal Gadot ou Scarlett Johansson (isso é regularmente sobre os rostos das mulheres) são usados para substituir os rostos originais (FLORIDI, 2018, p. 320).

O esforço desta prática reside na coleta de dados de feições humanas, expressões, traços faciais, por meio de vídeos e fotos originais, que são sobrepostos em outras imagens modificando o teor real das mensagens transmitidas a princípio. Esse tipo de técnica já vem sendo usado para criar notícias falsas e deturpar tanto no âmbito social como político, como aconteceu no caso Barack *deepfaking* Obama².

Em uma projeção para o futuro evidencia Chesney e Citron (2019, p. 1) que a popularização da *deepfake* potencializará a desinformação, principalmente no âmbito da política e relações internacionais, registrando que “podem ser usados para criar mentiras extraordinariamente eficazes capazes de incitar a violência”, razão pela qual esta deverá também ser considerada como uma modalidade de manifestação da desinformação.

A *deepfake* também é uma manifestação que envolve conjuntamente a desinformação como **processo** quanto como **propósito**. O *deepfake* envolve a dissimulação de se passar por alguém da forma mais crível possível. Em alguns filmes mais antigos de espionagem, observava-se que o agente em muitos casos utilizava máscaras realistas de um determinado alvo para enganar outrem. Hoje em dia, o surgimento da editoração eletrônica, em níveis cada vez mais realistas, possibilita que farsas bastante elaboradas envolvendo voz e vídeo sejam criadas de forma eletrônica e em uma trama (processo) elaborada com o intuito de se conquistar um objetivo

² Deepfake produzida envolvendo a imagens e vídeos do ex-presidente Barack Obama. Disponível em: < <https://www.vox.com/2018/4/18/17252410/jordan-peelee-obama-deepfake-buzzfeed>>. Acesso em: 12 fev. 2019.

(propósito), assim como nos filmes de espões, a *deepfake* se mostra como uma ferramenta.

A relação com a desinformação como **coisa** é mais forte, uma vez que a informação visual/auditiva faz parte do contexto onde a *deepfake* é realizada, e o autor da mentira sabe que a informação sensorial é prejudicada, e conseqüentemente, tem-se uma fonte de informação inverídica.

2.3 Movimentos Criados a partir da Desinformação

Com os constantes casos de desinformação aparecendo na grande mídia e com atos claros de prejudicar uma parcela maior da sociedade, pode-se observar o surgimento de alguns movimentos ou comportamentos criados a partir da desinformação. Alguns como combate a este fenômeno e outros como conseqüências dele. Iniciando por aqueles que vão de encontro a escalada da desinformação, é possível destacar o vigilantismo e a checagem de fatos, e como resultado deste processo, tem-se os filtros bolhas (bolhas) e a tecnocensura.

2.3.1 Vigilantismo

Conforme Johnston (1996, p. 226), “o vigilantismo surge quando alguma ordem estabelecida é percebida como ameaçada da transgressão (ou transgressão potencial) das normas institucionalizadas”, desta feita, é possível notar o porquê desse movimento ser desencadeado a partir do fenômeno da desinformação, posto que este desrespeita valores e garantias socialmente estabelecidos como o direito à informação.

Segundo Starbird *et al.* (2014), é o fenômeno em que as pessoas se sentem compelidas a combater a desinformação, seja por questões de crença ou até mesmo por senso de dever. Mascaro e Sperber (2009) apontam que esse é um comportamento parecido ao de pais que controlam o conteúdo que os filhos podem acessar na Internet. Cheong e Gong (2010, p. 474) entendem que a prática se torna mais comum “à medida que os participantes on-line se envolvem com a mídia e uns com os outros para divulgar informações e / ou criar textos”.

Hodiernamente há uma preocupação dos serviços de Internet, tais como Google, Facebook e Twitter para eles se tornarem os primeiros filtros para a desinformação (ZATTAR, 2017). Assim sendo, compreende-se o vigilantismo na Internet como um esforço formal (iniciativas de organizações e instituições públicas ou privadas) ou informal (iniciativas de particulares de maneira individual ou coletiva) que buscam fiscalizar, corrigir, alertar, viabilizar o bom uso e disseminação de informações na Internet.

Esses esforços tendem a se tornar cada vez mais comuns, ao passo que a Internet se desenvolveu como um ambiente não regulado do ponto de vista jurídico e social onde as pessoas podem ser controladas através de atos de vigilante. “O vigilantismo é frequentemente a resposta preferida, e às vezes a única eficaz, para fazer com que os ‘colonos cibernéticos’ percebam os crimes contra a propriedade e as pessoas” (MCLURE, 2000, p. 463).

Apesar da ideia de vigilantismo estar comumente associada a censura e regulação, para fins deste trabalho, atem-se à concepção do vigilantismo na Internet que visa “o controle social para corrigir e invocar um normal aspecto da vida social no mundo cibernético” (WEHMHOENER, 2010, p. 11) como forma de conter ou corrigir a desinformação.

Um exemplo recente que efetivamente resultou na prática de vigilantismo na Internet, ocorreu em ações realizadas no *Twitter* em defesa da vereadora Marielle Franco que foi alvo de uma série de notícias falsas nas redes sociais após o seu assassinato. Após a primeira onda de calúnias, se iniciaram as medidas de vigilância e a repressão aos boatos que tiveram tamanho engajamento da sociedade e repercussão, que as mensagens de esclarecimento superaram o volume de informações falsas sobre a parlamentar³.

2.3.2 Checagem de Fatos

³Disponível: <<http://dapp.fgv.br/reacao-boatos-superou-difusao-de-informacoes-contra-marielle-no-twitter-aponta-estudo-da-fgv-dapp/>>. Acesso em: 12 fev. 2019.

A checagem de fatos ou *fact check* surge como uma contrapartida as investidas resultantes da desinformação, principalmente no que diz respeito às notícias falsas, meias verdades ou pós-verdade. Para Mosseri (2016), essa prática é materializada através de ferramentas que permitem aos usuários da Internet sinalizarem e denunciarem conteúdos inverídicos ou mal-intencionados, surgindo, portanto, como um meio de combate e prevenção ao fenômeno da desinformação.

De acordo com Zattar (2017), essa verificação de fatos permite que o usuário mensure o grau de fidedignidade que possui uma informação a qual foi exposto. Verifica-se, portanto, que esse movimento fez surgir uma série de instituições cujo objetivo é prestar esse serviço de checagem de forma prática e desburocratizada, como por exemplo, o Lupa⁴, o PoliticFact⁵, o FactCheck.org⁶, o Chequeado⁷, o Truco⁸, AosFatos.org⁹.

Explica Zattar (2017) que:

As checagens de notícias procuram verificar uma informação a partir de classificações que, em geral, indicam os níveis de veracidade. A checagem pressupõe o uso de informações públicas e fontes confiáveis para verificação de conteúdo, o que resulta em avaliações que visam indicar o nível de veracidade de uma informação (ZATTAR, 2017, p. 289).

Iniciativas governamentais também despontam neste cenário com o objetivo de garantir a lisura de processos democráticos, como no caso das eleições, em 2018, em que o Tribunal Superior Eleitoral (TSE) tratou de criar um portal denominado “Esclarecimentos sobre informações falsas” onde qualquer eleitor pode buscar informações para esclarecer boatos e notícias com conteúdo de origem duvidosa disseminadas com o intuito de confundir-los.

Apesar destas ações, é possível notar que a maioria dos consumidores de informação na Internet não se valem das ferramentas de checagem no momento de disseminar o conteúdo recebido, seja por falta de conhecimento quanto ao manuseio ou existência desse tipo de serviço, seja por ausência de cuidado acarretada pela

⁴ Disponível em: <<http://piaui.folha.uol.com.br/lupa/>>. Acesso em: 12 fev. 2019.

⁵ Disponível em: <<http://www.politifact.com/>>. Acesso em: 12 fev. 2019.

⁶ Disponível em: <<http://factcheck.org/>>. Acesso em: 12 fev. 2019.

⁷ Disponível em: <<http://chequeado.com/>>. Acesso em: 12 fev. 2019.

⁸ Disponível em: <<http://apublica.org/truco/>>. Acesso em: 12 fev. 2019.

⁹ Disponível em: <<https://aosfatos.org/>>. Acesso em: 12 fev. 2019.

necessidade de replicar a informação. Para tanto, alerta Silva (2017) que em 2016 pelo menos 78% dos brasileiros utilizavam as redes sociais como meio para se informar, 42% afirmaram já ter sido alcançado por uma notícia falsa, mas que apenas 39% do total buscavam checar a veracidade e fonte do conteúdo informativo.

Por essa razão, mesmo com a checagem de fatos se apresentando como um movimento decorrente da onda de desinformação vivenciada na Internet, a sua eficácia enquanto alternativa para o combate e prevenção deste fenômeno restará reduzida, posto a ausência de consciência e cultura dos usuários da rede quanto a importância da checagem, sendo a disseminação em massa uma das propulsoras desse processo.

2.3.3 Tecnocensura

Mídias sociais como *Twitter*, *Facebook*, *Weibo*¹⁰ e *Whatsapp* surgem no contexto da Internet como meios de promoção e organização de movimentos populares que visam a reivindicação de pautas políticas ou sociais, evidenciando aparecimento de um ativismo social mais latente nesses ambientes, que conseqüentemente passam a ser espaços de promoção para a censura tecnológica.

Diamond (2010) reconhece essas tecnologias como um meio de promoção da liberdade de expressão e manifestação do cidadão que “permite aos cidadãos denunciar notícias, expor irregularidades, expressar mobilizar protestos, monitorar eleições, fiscalizar o governo, aprofundar a participação e ampliar os horizontes da liberdade” (DIAMOND, 2010, p. 70).

Entretanto, Trottier (2014) alerta para as conseqüências e reações a essas tecnologias, pontuando “o uso dessas mídias durante recentes protestos no Oriente Médio e na China, regimes nessas regiões fizeram esforços elaborados e agressivos para controlar a Internet comunicação” (TROTTIER, 2014, p, 189). Mesmo entendimento expressa Shirky (2011, p. 29) relatando que “as mídias sociais se tornaram ferramentas de coordenação para os movimentos políticos, assim como a maioria dos governos autoritários do mundo está tentando limitar acesso a ele”.

¹⁰ Twitter Chinês

Aparece, assim, o termo tecnocensura (ou tecnocensória) que, como bem definiu Baker (2015), é a prática de silenciar vozes na Internet, referindo-se ao momento político ocorrido no México em 2012, onde a promoção de desinformação em massa se deu através do emprego da *Data Flood* e *Social Spam* no *Twitter* como forma de abafar protestos populares contra o então presidente Peña Neto.

Para Trottier (2014) a censura é caracterizada basicamente pela vigilância do ambiente *online*, entretanto não se deve confundi-la com o vigilantismo retratado aqui, mas sim com o intuito de controlar as ações de protestos, manifestações, organizações e ter acesso a informações dos usuários da Internet, como os programas de vigilância dos EUA denunciados por Edward Snowden¹¹.

Exemplificando o que denota Trottier (2014), verifica-se que países como o Irã, Tunísia e China utilizam deliberadamente da tecnocensura para controlar a comunicação na Internet. Para o autor, esses países exercem controle seja pelo bloqueio de sistemas de informação (nacionais ou internacionais), seja pela filtragem de mensagens desenvolvendo assim métodos de direcionamento para censura.

Desta forma, reconhece Trottier (2014) que as práticas de censura na Internet têm evoluído, atribuindo esse fator ao direcionamento de conteúdo aos usuários a partir de estudos de perfis e bloqueio de conteúdo promovido pelas plataformas ou sites, se alinhando a interesses do Estado ou das Organizações públicas e privadas.

O autor relata que uma das armas para exercício da tecnocensura utilizada por Estados como o Irã e a China, por exemplo, é a promoção de “um clima de medo e incerteza em relação à informação on-line, uma vez que número de sites são filtrados e bloqueados, e a desinformação é espalhada através de falsos sites de oposição para enganar ativistas e sufocar a dissidência” (TROTTIER, 2014, p. 193).

Outro ponto explorado para promoção da censura na Internet é a contratação de usuários para promoção de mensagens, através do anonimato, como forma de “orientar a direção da discussão pública de acordo com os princípios do Estado departamentos de propaganda” (TROTTIER, 2014, p. 196). Sendo assim, nota-se que a tecnocensura pode ser promovida tanto pela desinformação em suas diferentes modalidades de

¹¹ <https://exame.abril.com.br/tecnologia/o-mundo-sob-vigilancia-veja-cronologia-do-caso-snowden/>

apresentação, observando-se para tanto o seu intuito de conter as manifestações na Internet.

2.3.4 Filtro Bolha

Em 2011, Pariser cunhou o termo “*Filter Bubble*” que, traduzido para o português, significa Filtros Bolha. De acordo com o autor, essa prática consiste em mecanismos que criam um universo de informações na Internet de maneira direcionada ao usuário a partir das suas preferências e históricos de busca. Assim relata que “são mecanismos de previsão que criam e refinam constantemente uma teoria sobre quem somos e sobre o que vamos fazer ou desejar a seguir” (PARISER, 2011, p. 14) alterando, portanto, as informações e ideias dispostas em rede conforme o perfil do usuário.

A promoção dessas bolhas é realizada por máquinas que utilizam os rastros de usuários da Internet para definir suas preferências e promover a filtragem como meio de otimizar o processo de busca. Como observado, “os algoritmos estão sempre interpretando ações de pessoas conectadas, através de inúmeras variáveis, para propor uma experiência mais próxima do que essas máquinas acreditam ser o que se está buscando” (FERREIRA; RIOS, 2017, p. 2).

Deste modo, nota-se que esses filtros podem surgir tanto em um ambiente de sobrecarga informacional, como ocorre no caso do *Data Flood* ou *Social Spam*, quando nos demais processos de desinformação promovidos através de outras manifestações. Magrani (2014, p. 118) destaca que esse mecanismo pode ser entendido como “conjunto de dados gerado por todos os mecanismos algorítmicos utilizados para se fazer uma edição invisível voltada à customização da navegação on-line”.

Ou seja, trata-se de uma personalização de conteúdo na Internet que determinado usuário receberá ou terá acesso a partir do perfil traçado por meio dos filtros estabelecidos, sejam estes promovidos de forma autônoma pelo provedor, seja ele de forma voluntária pelo usuário. Assim seria:

A filtragem de conteúdo feita pelos próprios provedores, limitando nossa autonomia on-line na busca por novas informações na esfera pública virtual. De outro lado, há a busca dos indivíduos por filtrarem voluntariamente as informações que consomem gerando o problema da fragmentação do debate a

partir do momento em que as pessoas vão optar por consumir informações que estejam alinhadas com seus próprios pontos de vista (MAGRANI, 2014, p. 117).

A ocorrência dos filtros-bolha representa uma ameaça aos meios de combate e prevenção à desinformação, promovendo a polarização das discussões, pois limitam o conteúdo recebido pelo usuário conforme suas convicções, gerando cada vez mais desinformação daquele que foi afetado por esse processo, já que os usuários tendem a se afastar dos pontos de vistas distintos a partir de uma definição algorítmica.

Os filtros-bolhas ocasionam cortinas de fumaça que surgem como forma de afogar debates, manifestações ou reivindicações e autonomia dos indivíduos que, além de serem induzidos por robôs que atuam em rede, passam a cooperar inconscientemente para o fomento de conteúdos de origem duvidosa na Internet.

2.4 Síntese da Seção

Ao término desta seção, é possível observar que algumas percepções dadas pelo senso comum sobre desinformação não são exatamente o *status quo* sobre a mesma. Inicialmente, se tem a ideia de que a desinformação é fruto de um processo ardiloso, minucioso, intencional e com a finalidade do engodo e destrutiva. A partir do momento que se consideram as defesas naturais de insetos, com a intenção de sobreviver, desinformação, é plausível imaginar que a desinformação está em um grande eixo onde há extremos entre destruir-sobreviver, enganar-tranquilizar, simples-complexo, intencional-não intencional, mentira-verdade e tantos outros objetos antagônicos que fazem parte deste fenômeno maior.

Então é a partir destas dualidades, e de toda zona cinza que existe entre seus extremos, que foi trazido o objeto desinformação e que todo esse contexto só atribui a ela a complexidade e amplitude das ações cotidianas de desinformar. Ao caracterizar a desinformação como um fenômeno social, encontra-se diversas situações em que os indivíduos tentam: “contar vantagem”, “sair por cima”, “criar boa impressão”, “relevar situações adversas”, “engolir sapos”, “sorriso amarelo”, entre tantas outras. Pessoalmente, coube até a reflexão de que a desinformação é algo tão frequente ou mais do que a informação no mar de omissões e meias verdades presentes no dia-a-dia.

Por ser um objeto tão vasto e rico, foi escolhido como um ponto de partida discutir três das facetas da desinformação (coisa, processo e propósito). Resumidamente foram apresentadas características e finalidades de cada uma dessas facetas, conforme o quadro a seguir:

Quadro 2 - Resumo das Facetas da Desinformação

Desinformação como Coisa	Desinformação com Processo	Desinformação como Propósito
Falhas ou imperfeições no objeto informação	Desinformação é o resultado de um defeito no processo de informar	Diminuição da capacidade de reflexão dos indivíduos
Falha do registro da informação em seu suporte	Execução de ações sistematizadas (processo) para se atingir o objetivo de desinformar	Implementação de projetos de poder
Falha na transmissão da informação		Impossibilidade de acesso à informação de qualidade por parte da maioria das pessoas
Falha na percepção/compreensão/uso da informação por parte do interagente		

Fonte: Autoria Própria

Apesar de aparentemente bem divididas quanto as suas caracterizações, sem dúvidas, existem momentos em que essas facetas se misturam, aparecem de forma conjunta e até mesmo simbióticas (uma não existem sem outras), mas, espera-se ter esclarecido que estas três facetas podem revelar olhares totalmente distintos e complementares de todo o panorama.

Em seguida, foram apresentadas uma série manifestações que, embora não sejam recentes, têm ganho espaço na última década pela grande mídia e parece se materializar em um espaço de grande uso social (Internet) e que não vem sendo discutido com tanto vigor na Ciência da Informação. Portanto, também se considera a seção de manifestação como um ponto de partida, pelo menos para algumas delas, e observá-las de uma maneira mais global e menos pontual. Certamente as notícias falsas e pós-

verdades são objetos mais discutidos na CI, mas nem sempre estes estão em um contexto de desinformação ou relacionados.

Por fim, essa percepção da desinformação por parte da sociedade deveria apresentar algum tipo de reação. E foram apresentados alguns movimentos que decorrem a partir desta desinformação em vigor. Os primeiros são reações da sociedade contra a desinformação numa tentativa de tornar os ambientes informacionais isentos. Os últimos, são questões relativas à alienação por meio da criação, intencional as vezes, de bolhas e da censura promovida por entidades de poder. Esta seção também é um ponto de partida, e não se pretende, de forma alguma limitar a estes movimentos.

Ou seja, considera-se que esta seção seja um eterno “em construção” esperando novas colaborações e visões. Todavia, que já é suficiente para lastrear uma visão sobre como máquinas sociais podem apoiar esse fenômeno.

3 MÁQUINAS SOCIAIS E ENTIDADES NÃO BIOLÓGICAS

Máquinas Sociais é fenômeno relativamente recente, cujos debates acerca do tema datam do início da década de 2010; se trata de um conjunto de *software* que está presente na Internet contemporânea e que tem como princípio atuar em cima dos dados sociais gerados na rede com o intuito de prestar serviços para as pessoas. Entretanto, antes de iniciar a reflexão sobre Máquinas Sociais, faz-se necessário posicioná-las no cenário que envolve entidades não humanas em um contexto geral.

A necessidade dessa explanação se dá pela diversidade de contextos em que alguns destes atores não humanos são mencionados de forma sobreposta, confusa e até mesmo enganosa. Por isso, para melhor posicionar as máquinas sociais, é imprescindível definir alguns destes atores, tais como Robôs, Máquinas, dentre outros, bem como, contextualizar as próprias Máquinas Sociais e assim justificar o porquê da escolha deste tipo de *software*.

Desta forma, será iniciada a explanação desses “conceitos” pelos robôs e a descaracterização deles como exclusivamente de entidades não humanas.

3.1 Robôs

Ballard (2014) aponta que o termo “robô” tem a sua primeira aparição na literatura no ano de 1921, sendo introduzido pelo romancista Karel Capek na obra *Rossumovi Univerzální Roboti* oriundo das palavras checas “*robotnik*” que significa trabalhador forçado ou da palavra “*robotá*” ou “trabalho forçado”. Segundo o escritor, um robô seria “uma entidade inteligente e de aparência similar aos seres humanos, cujo papel era tornar a vida das pessoas mais agradável uma vez que esta realizava todo o tipo de trabalho que os humanos não gostam de fazer” (BALLARD, 2014, p. 356).

Na obra original, os robôs não eram mecânicos nem tão diferentes dos humanos. Eles eram seres biológicos criados para trabalhar para os humanos, sendo mais ou menos indistinguíveis dos homens na aparência e até mesmo com habilidades cognitivas completas. Na ficção, os robôs se tornam onipresentes, assumindo praticamente todo

trabalho da humanidade e, em certa altura, eles se cansam de fazer todo o trabalho para os “humanos preguiçosos”.

Santana e Lima (2018) também resgatam a visão ficcional do que seria um robô ao destacar a dualidade entre estes e os humanos representada pelos personagens do conto o Mágico de Oz, datado do início do século XX, trazido originalmente pela versão como o homem de lata. Este foi chamado pelo próprio autor da obra (Eduard Baum) de “Protótipo Robô”. No trabalho, os autores fazem a releitura de que na visão mais contemporânea dos robôs, o espantalho (cérebro), e não o homem de lata (coração), seria o ator principal em uma possível relação de oposição entre homem / robô.

Outras características referentes a estas entidades vão além do serviço a humanidade. Oliveira (2010) observa a questão da programabilidade e autonomia de tais entes. O autor sugere que os robôs são entidades “programadas para executar movimentos fixos, sem que haja necessidade de se monitorar, ou ainda representar, o ambiente na qual estão inseridos” (OLIVEIRA, 2010, p. 22), verifica-se, portanto, que os robôs são uma espécie de tecnologia voltada para realização de tarefas repetitivas, de média ou baixa complexidade, que anteriormente era executada por um humano.

Ainda que a questão da programação e autonomia sejam destacadas no trabalho de Oliveira, pode-se observar que o autor não dissocia estas características de um corpo tangível, afinal, ainda se condiciona ao robô a necessidade de movimentos fixos, o que sugere a existência de uma estrutura física no contexto que envolve esta entidade.

A visão de que um robô está vinculado a algum tipo de forma (corpo) foi superada com o passar do tempo. Os robôs, que são criados com o intuito de ter uma forma, foram redenominados andróides (categoria que será explanada mais à frente). Desta forma, se abre a possibilidade para que algumas definições de robôs sequer tenham algum tipo de corpo/*hardware* tangível. Estes trabalhos que dissociam os robôs de corpos físicos começaram a ser publicados no início dos anos 2000.

Por exemplo, na Ciência da Informação se pode apontar o trabalho de Pierozzi Júnior *et al.* (2003) no qual robôs são classificados como um gênero da espécie agentes, sendo “um programa que vasculha automaticamente a estrutura de hipertexto da Internet procurando por um arquivo e continua buscando recursivamente todos os arquivos que são referenciados nessa estrutura” (PIEROZZI JÚNIOR *et al.*, 2003, p. 105). Portanto,

esta classificação trazida pelos autores põe os robôs como uma subcategoria do termo agente e que realiza uma atividade específica de busca e recuperação de informação na Internet.

Ou seja, aqui os autores tratam os robôs como ferramentas que atuam na busca e recuperação da informação na rede, de modo que alimenta as bases de dados que servem como acervo para os buscadores como o *Google*, por exemplo. Entretanto, será visto mais à frente que esta classificação está mais adequada ao que conhecemos como *bots*.

Observa-se que, mesmo propondo um novo olhar do que seriam os robôs, talvez o termo não seja o mais adequado para representar a entidade apontada pelos autores, mas a analogia de sistemas de *software* com robôs é compreensível. Vale ressaltar que não é exclusividade destes autores o uso, discutível, do termo robôs para tais tipos de *software* e a confusão do uso deste com o termo *bot* e agente.

Seguindo outra linha, alguns pesquisadores começam a se debruçar sobre comportamentos humanos que agem de forma descuidada ou que possuem menos interesses em atividades complexas e que agem como robôs devido a simplicidade cognitiva em suas ações. Carr (2011) é um dos percussores nestes estudos e, em momentos de sua obra, ele compara alguns comportamentos desta nova geração à robôs.

Assim é possível concluir que a principal característica dos robôs é, em seu primeiro momento, a subserviência aos seres humanos para realizar atividades indesejadas/não recomendadas. Também se faz presente nestes atores a questão de uma maior autonomia (não supervisão humana) e da necessidade de um componente físico (*hardware*). Não é atribuída aos robôs a obrigatoriedade de uma natureza mecanizada. Apenas ao longo dos anos se foi atribuído a estes o caráter eletromecânico e a relação destes com a programação.

O resgate original deste conceito é importante para a Ciência da Informação, uma vez que se pode evitar, não só a clara confusão do uso da terminologia, mas também, reabrir as possibilidades de estudos de comportamentos robóticos em seres biológicos, inclusive humanos, determinando que a característica mais importante de um robô não é a automação ou mecanização, mas sim a ausência de anseios próprios sendo os robôs

“escravos” de outra entidade, que não é necessariamente mais inteligente ou que sequer tenha algum tipo de consciência e cognição. Um exemplo disto é a área Computação Humana que investiga como os seres humanos podem auxiliar os sistemas digitais a resolver os seus problemas. Neste cenário, a cognição humana é usada como ferramenta por um computador para solução dos problemas.

Isso significa que o uso do termo robô para se denotar a todo e qualquer tipo de entidade não biológica (incluindo *software*), além de excluir os seres vivos deste universo dos candidatos a robôs, não é o mais adequado. Isso significa que o termo robô em si é utilizado algumas vezes de forma equivocada. Entretanto, se robôs são entidades que podem ser tanto biológicas quanto eletromecânicas, seria o termo máquina, então, o mais adequado para se referir a estes seres artificialmente concebidos?

3.2 Máquinas

O termo *machina* aparece no latim desde Plauto e Ênio, no início do século II a.C. e seu uso se tornou cada vez mais comum durante a era imperial e a antiguidade tardia (século VIII d.C). Esta palavra foi emprestada do vocabulário Dórico e tem como significado “meios, artifícios e dispositivo” e não distingue entre meios materiais e imateriais, permitindo que estes se sobreponham ou que se misturem (RAUNIG, 2008).

A primeira vez que uma máquina foi apresentada como um ser mais próximo a um robô foi no filme alemão *Metrópolis* de 1927 ao apresentar o personagem Máquina-Homem (*Maschinenmensch*). Mas a visão de que estruturas mecânicas complexas podem ser chamadas de máquina remonta ao vocábulo original da palavra, o qual sugere que uma máquina se refira a uma estrutura (seja um objeto ou organização) que realiza uma determinada tarefa ou que produza um resultado específico, de modo a simplificar, ou até mesmo abstrair, como a ação foi realizada. Raunig (2008) ressalta esta definição ao explicar a expressão “*Deus Ex Machina*” que se refere aos desfechos teatrais gregos do século II d.C que ocorrem de forma súbita. Nesse caso, o termo denotava situações no palco onde seres voadores (normalmente deuses) eram suspensos por estruturas muito específicas, chamadas máquina, e promoviam o desfecho do espetáculo.

O uso da nomenclatura máquina para se referir às estruturas complexas também foi seguido na concepção dos primeiros dispositivos para processar informação. No século XVII, surgiu a calculadora Pascalina, criada por Blaise Pascal em 1642, que ficou conhecida popularmente como a máquina aritmética. A palavra foi sacramentada por Alan Turing quando este publicou a ideia da Máquina Universal de Turing na década de 1930, que é a base dos computadores modernos que usamos hoje em dia.

A aplicabilidade do termo máquinas para denotar estruturas, ao invés de dispositivos mecânicos, pode ser observada na visão particular de Newton, que acreditava que o mundo era uma máquina criada para servir ao seu Mestre. Esta visão, segundo Maani (2005) se assemelha muito a “Metáfora da Máquina” cunhada por Gareth Morgan, onde uma ideia estritamente mecanicista é atribuída às organizações. Nesta concepção, as instituições são divididas em elementos (máquinas) que produzem seus resultados de forma individual contribuindo para o sucesso coletivo. Segundo Carlson e Gjersvik (1997), o foco desta metáfora está no controle, repetição, previsibilidade, trabalho especializado de partes substituíveis realizados por qualquer sistema ou organização.

Na Ciência da Informação o vocábulo ‘máquina’ possui uma utilização bastante abrangente, indo desde a representação de dispositivos mecânicos/eletrônicos até a sugestão de estruturas sociais. Ressalta-se que a visão de interesse para este trabalho é a relação entre o termo máquina e a representação de dispositivos concebidos para propósitos de processamento de informação. Neste sentido, é comum encontrar a atribuição desta nomenclatura a dispositivos eletrônicos seguindo a do próprio Alan Turing. A expressão máquina, assim como o termo robô, possui algumas características peculiares e que são distintas de dispositivos mecânicos/eletrônicos.

Entretanto, é importante chamar a atenção para o fato que o significado desta palavra foi modificando ao longo do tempo. Inicialmente o conceito máquina denotava uma estrutura complexa (seja um dispositivo ou organização) que realizava uma determinada transformação (de produtos ou processos) de forma abstrata (caixa-preta). Ou seja, isso significa que a metáfora máquina remete a uma realidade onde determinados problemas eram resolvidos por uma entidade cujo funcionamento e

mecanismo de ação não precisavam ser conhecidos pelas pessoas que utilizavam/operavam tal máquina.

Então é bastante compreensível que o uso deste termo se aplique a dispositivos complexos cujo funcionamento não precisa ser totalmente compreendido pelos usuários. Ao imaginar um carro, por exemplo, um motorista não precisa compreender o sistema elétrico ou de força para dirigir. Toda a complexidade da estrutura do carro é minimizada por representações dadas por objetos (volante, pedais, marcha e botões) que são uma espécie de interface para o uso. Também é compreensível a adoção desta locução para estruturas complexas, seja de ordem da magnitude do tamanho de galáxias (Newton) ou de organizações para observar os fenômenos de interesse e se abstrair os detalhes que não são interessantes.

Esta ideia de que a palavra 'máquina' se aplica tanto a dispositivos quanto a estruturas encontrou visões comuns, que se aplicam aos dois contextos, de que estes poderiam ser compostos por partes menores, que sempre entregavam resultados padronizados, e, portanto, previsíveis, em processos que poderiam ser executados repetidas vezes e onde parâmetros podem ser definidos para atingir determinados níveis (controle).

Na pesquisa, não foram encontrados trabalhos na área da Ciência da Informação que trouxessem uma definição de máquina e por isso se fez necessário remeter ao contexto histórico deste conceito e como este evoluiu ao longo dos séculos. O uso comum do termo na área se refere geralmente aos complexos dispositivos eletromecânicos ou a superestruturas que abrangem diferentes contextos, mas que, sem exceção, tem o poder para decidir, ou controlar, ou manipular situações em benefício próprio.

Desta forma, se os robôs têm como principal característica a subserviência e autonomia, as máquinas remetem às estruturas de abstração onde o observador/utilizador não necessita conhecer de forma detalhada o seu funcionamento. Entretanto, também se observa que as máquinas também trazem a ideia de partes substituíveis e que o produto de saída desta deve ser padronizado e obedecer a parâmetros previamente definidos.

Assim como os robôs, as máquinas também não se mostram como a representação mais adequada para se referir a entidades não biológicas. Desta forma, foi buscado diretamente um termo que se tratasse de tais entidades, assim chegando ao substantivo autômato.

3.3 Autômatos

Se máquinas e robôs não se referem especificamente a entidades que não são dotadas de “vida”, cabe aos autômatos representar esta classe de organismos artificiais, mas que realizam atividades de modo autônomo e independente. A palavra autômato é a forma latina da palavra grega *αὐτόματον* ou aquele que possui vontade própria. Bailly (2003) afirma que este termo foi utilizado pela primeira vez no livro *Ilíada* de Homero ao se referir a um dispositivo automático para a abertura de portas. Ainda segundo Bailly, a palavra foi evoluindo ao longo dos séculos e na Europa passou a tratar os dispositivos autônomos como aqueles que remetiam a ações de seres vivos, tais como o cuco ou os soldados de chumbo que martelavam os relógios.

Autômatos já eram figuras presentes na história humana desde a antiguidade como, por exemplo, a mulher forjada de ouro contida no poema Finlandês *kalévala* que, em meados do século XIX, destinava-se a tradições orais que datam de 1000 a.C (EDUARD, 1990). A mitologia helenística também apresenta os seus autômatos, tais como o homem de bronze artificial Talos, ou Dédalo, criador das estátuas que se movem (THEOI, 2000).

Choi (2011) faz uma reflexão da transformação das “máquinas” que estão ao nosso redor, as quais são cada vez menos programadas e cada vez mais independentes. Dispositivos eletrônicos estão progressivamente ganhando consciência e inclusive de sua própria condição de existência. Robôs, como conhecemos hoje, estão perdendo a simplicidade de ser subservientes e passam a se tornar cada vez mais conscientes de suas ações e de si mesmo e, segundo o autor, os autômatos deixam de ter como principal característica um corpo maquínico autônomo, para se tornar um ente mais adaptável e envolvido em “seus próprios pensamentos”.

A ideia de que autômatos não precisam de um “corpo” é amplamente observada na matemática, mais especificamente na teoria dos autômatos, que trata de máquinas abstratas. Chakraborty, Saxena e Katti (2011) afirmam que esta teoria se refere ao funcionamento de “máquinas”, puramente baseada em seus estados, por isso são consideradas abstratas. Não importam, nestes casos, que dispositivos ou mecanismos compõem tais máquinas e sim o seu resultado/funcionalidade. Essa imagem é a percussora do que hoje conhecemos como *Software*, onde máquinas de propósitos gerais (*Hardware*) seriam orientadas/programadas a partir de seus possíveis estados (*Software*).

Neste sentido, os autômatos trazem à tona dois elementos importantes para a criação desse panorama sobre entidades não humanas: (i) o primeiro é que, diferentemente de máquinas e robôs, o termo autômato são seres, pelo menos em sua gênese, não dotados de vida que realizam suas atividades com autonomia, seguindo sua vontade própria; (ii) o segundo é que esta vontade própria se caracteriza cada vez mais pela criação/construção de elementos voltados a decisão/automação concretizados hoje pelo que se chama de *Software*.

Assim, Choi (2011) afirma que a relação hodierna entre robôs e autômatos se dá no sentido de que o segundo é uma evolução do primeiro, ou seja, que o autômato surge quando são dados aos robôs a capacidade de tomar decisões que envolvam criatividade e que estão totalmente fora de um determinado padrão de programação. A pergunta feita, por Turing (1950), “Pode uma máquina pensar?”, revela uma inquietação de parte da comunidade científica que se debruça sobre o problema mente-corpo e que, segundo a teoria dos autômatos, robôs e máquinas podem ser dotados de “uma espécie de consciência” que os proveriam de uma forma diferente de vontade.

Não foram encontrados trabalhos na Ciência da Informação que se utilize do termo autômato, o que sugere um uso muito pequeno deste por pesquisadores da área ao ponto de determinar seu emprego. Entretanto, existe outro termo de uso comum para entidades não humana, mas que a função destes é se parecer com os humanos, são os chamados Androides.

3.4 Androides

Os elementos investigados até aqui apresentam características distintas sobre sua razão de ser, entretanto, nenhuma delas tem como finalidade a semelhança com estruturas vivas, tais como seres humanos e animais. Para esse objetivo, foi identificado o termo androide. A palavra se deriva do grego (ανδροειδής) e significa “aquele que se parece com um homem” e se atribui ao primeiro uso desta na língua inglesa ao alquimista Albertus Magnus, em 1727 (LAGRANDEUR, 2012). O uso da palavra foi atribuído a seres artificiais, que normalmente utilizam materiais que se parecem com o corpo humano.

Essa visão é corroborada por Regis (2006, p. 9) ao afirmar que os androides “se referem aos robôs com aparência humana, podendo ser produzidos com substâncias orgânicas ou revestidos com materiais sintéticos que imitam fielmente musculatura e pele”. Nóbrega (2016, p. 3) também destaca estas características nos androides ao ressaltar que “além de possuir a forma física humana, tal qual o humanoide, possui também a aparência semelhante à orgânica, sendo distinguível de seus criadores somente por testes específicos”.

O que se destaca na definição de Nóbrega (2016) é a diferença entre um androide e humanoide, já que, este último, pode até possuir forma humana, mas não há a necessidade de simular a aparência humana. Stableford (2016) aponta que se tornou comum o uso da palavra Androide para se referir a robôs humanoides.

Ishiguro (2005) relaciona os androides com os autômatos ao ponderar que existe a visão de que a morfologia humana não é, necessariamente, a única característica que deve observada na semelhança entre androides e humanos. O autor aponta que a capacidade em imitar os seres humanos é que denotam esses novos androides. Segundo o autor, enquanto os autômatos são “livres” e autônomos para agir de forma independente, existe uma nova categoria de autômatos, chamados de simulacros (do latim *simulacra*) que, no comportamento, mais do que na fisiologia, têm a intenção de imitar seres humanos.

No trabalho de Turing (1950), ele propõe um desafio conhecido como Jogo da Imitação, também conhecido como Teste de Turing, no qual o objetivo da “máquina” seria responder perguntas como se fosse um ser humano, com intuito de enganar outra

pessoa. Então se abre a possibilidade de que Androides possam ser enxergados como “aquele que se parece com um homem” também em sua forma de “pensar”. Tais simulacros já existem e são encontrados na Internet e foi classificado por de Oliveira, Vieira e Lopes (2016) como *Socialbots*.

Assim é possível observar que as características determinantes dos androides é a capacidade em se parecer (imitar) seres humanos. Se em um primeiro momento histórico isso vinha da fisiologia e da busca pela perfeição em recriar fielmente as características humanas, hoje a busca pela semelhança com os seres humanos tem um viés comportamental, uma vez que as interações virtuais têm se tornado a maneira mais usual de se comunicar. A partir da expansão do uso da Internet onde simulacros, sem corpos, atuam como seres humanos, torna-se necessário esse olhar sobre os androides, ainda pouco explorado, em que se observa a imitação da “mente” e não do corpo.

E o que acontece quando seres vivos são “artificializados”? Ou seja, do que se pode chamar as entidades biológicas que se misturam com partes eletromecânicas para formar um ser que não é completamente natural, mas que não são autômatos ou humanoides? A resposta para esta pergunta são os ciborgues.

3.5 Ciborgues

A entidade resultante da junção de organismos biológicos e eletromecânicos em um mesmo contexto são os ciborgues. A criatura referente a este questionamento pode não ser representada pelas quatro anteriores apresentadas até aqui. O termo ciborgue é utilizado quando se quer destacar a hibridização entre seres vivos e partes mecânicas, surgindo como aquele que representa esta associação entre homens e “máquinas”.

De acordo com Andrade (2008), a palavra “*cyborg*” foi cunhada em 1960 por Manfred Clynes e Nathan kline, cujo significado seria “homem ampliado”. A origem do termo deriva da junção das palavras **cib**ernético e **organ**ismo, traduzindo a ideia de que este se refere “a um ser vivo com partes do corpo orgânicas e biomecânicas auxiliadas ou controladas por dispositivos mecânicos ou eletrônicos” (NING *et al.*, 2017, p. 783). Santos *et al.* (2017, p. 81) corroboram essa opinião ao afirmar que ciborgues se tratam

de “redes entrelaçadas que são partes humanas e partes “máquinas”, nas quais a fronteira entre o humano e a máquina são redefinidas”.

Ambos se aproximam da concepção trazida por Haddow *et al.* (2015) que descrevem essa entidade como um híbrido formado pela relação máquina-homem, cuja descrição é feita por Park (2014, p. 304) como “um ser humano com um dispositivo eletrônico implantado ou permanentemente ligado ao seu corpo com a finalidade de melhorar seus sentidos ou habilidades individuais”. Neste sentido, é possível identificar uma semelhança com a concepção de “homem ampliado” que surgiu na década de 1960, onde a tecnologia passa a ampliar as habilidades humanas tornando por vezes mais ágil, eficaz, eficiente ou precisa.

Quando tratamos da temática na área da Ciência da Informação uma das definições mais utilizadas para os estudos de Ciborgues pertence a Haraway (2009), o qual afirma que:

Um ciborgue é um organismo cibernético, um híbrido de máquina e organismo, uma criatura de realidade social e também uma criatura de ficção. Realidade social significa relações sociais vividas, significa nossa construção política mais importante, significa uma ficção capaz de mudar o mundo (HARAWAY, 2009, p. 36-37).

Entretanto, outros pontos de vista começam a surgir ampliando a ideia para além da simplória relação máquina-homem, fomentando novos horizontes para tratar do tema. Destaca-se nesse contexto o trabalho de Santaella (2004, p. 33), que afirma que os ciborgues são todos aqueles que estão “em estado de prontidão conectando-se entre nós e nexos, num roteiro multissequencial e labiríntico que ele próprio ajudou a construir ao interagir com os nós entre palavras, imagens, documentação, músicas, vídeos, e outros”, enquanto Chu *et al.* (2012, p. 814) afirmam que os “ciborgues são *bots* assistidos por humanos ou humanos assistidos por *bots*”.

Isso significa que essa hibridização entre seres vivos e componentes artificiais não se dá apenas no contexto físico, mas também em uma visão sistêmica. Porterfield (2018) aponta que hoje existe uma “orquestração” entre as atividades humanas e de *bots* na Internet de forma coordenada, colaborativa ao ponto que, se usa-se o termo ciborgue para se referir a seres humanos que tiveram suas capacidades ampliadas por

dispositivos eletromecânicos, por que não se pode considerar um ciborgue uma pessoa que se utiliza de algum tipo de *software* para potencializar seus limites cognitivos?

Nesta visão, a junção homem-máquina também não se dá por uma questão fisiológica, mas pela harmonização entre mente e capacidade computacional para realizar problemas que nem o homem, nem a máquina poderiam resolver sozinhos. Essa visão sistêmica de que máquinas estendem a capacidade humana para processar informação é estabelecida por Van Dijck *et al.* (2018) como “A sociedade da plataforma” onde pessoas e *software* vivem em um estado de simbiose.

Assim, conclui-se que a *Web* 3.0 também passa a transformar o conceito de Ciborgue, como foi com os termos Androide e Autômato. A partir do momento em que ecossistemas de pessoas e *software* também passam a ser considerados como os novos ciborgues, é possível compreender que a visão original do termo, onde humanos e “máquinas” são posicionados em um mesmo contexto (e não mais corpo) para aumentar as suas capacidades (seja do homem, seja da máquina), continua contemporâneo.

Mas existem entidades que não são baseadas em sistemas físicos? Ou seja, será que todo autômato precisa necessariamente de um componente material ou um corpo? Hodiernamente, são encontradas entidades de *software* que independem de *hardware* para sua existência. Começando pelos agentes de *software*.

3.6 Agentes de *Software*

Até aqui, foram destacadas as principais categorias em que os atores não humanos são normalmente classificados. Entretanto, o que se ressalta no senso comum sobre estes elementos é que, em sua maioria, estão atrelados a algum tipo de corpo (*hardware*). Mas, e aquelas entidades que não se utilizam de um maquinário tangível para realizar as suas atividades? Aqui se pode citar o *software* (categoria abrangente), entretanto, o *software* está para os elementos digitais assim como objetos (categoria abrangente) está para os elementos concretos. A partir desta ótica foi selecionada, dentro deste universo, uma categoria que represente aqueles “programas de computador” que possam agir de modo similar com as categorias vistas até aqui. E o ponto de partida se deu a partir dos agentes de *software*.

Segundo Bhargava (2019), um agente de *software* é um programa de computador que atua representando um usuário (humano ou outro agente) em um relacionamento de delegação. Ou seja, o agente tem a permissão de agir em nome do proprietário e a tal "delegação" implica em autorização sobre a decisão de quais ações são apropriadas a serem tomadas.

Nwana (1996) aponta que a definição de agente de *software* é abrangente e pode se referir a programas que utilizem, ou não, *hardware* de qualquer espécie (inclusive nenhum), pode ser autônomo ou trabalhar sob a supervisão de outras pessoas e agentes, pode ser supervisionado ou não e possuir qualidades humanas, tais como compreender a linguagem natural. Mas, o que o diferencia de um *software* qualquer é que o mesmo não é expresso a partir de suas funções e sim pelo seu comportamento.

Isso quer dizer que o agente não é concebido para realizar tarefas específicas e sim para cumprir metas mais amplas, podendo ser realizada uma tarefa ou mais, em ordens diferentes, executadas de formas diferentes, por atores distintos, entre outras características que estas tarefas podem variar.

Mahmoud (2000) classifica os agentes de *software* nos seguintes tipos:

- **Agentes Inteligentes:** Agentes que utilizam algum padrão de inteligência artificial para melhorar seu comportamento;
- **Agentes Autônomos:** Agentes capazes de modificar as suas estratégias para atingir os objetivos;
- **Agentes Distribuídos:** Agentes capazes de se "localizar" fisicamente, em vários computadores diferentes;
- **Agentes Colaborativos:** Agentes que são concebidos para trabalhar em conjunto com outros agentes;
- **Agentes Móveis:** Agentes que podem modificar o local físico de onde estão sendo executados;
- **Agentes de Interface:** Agentes que tem como principal objetivo auxiliar usuários (normalmente humanos) a utilizar algum sistema;
- **Agentes de Informação:** Agentes especializados em coletar, gerenciar e manipular informação oriunda de diferentes fontes.

Nwana (1996) afirma que estes agentes possuem características particulares:

- **Trabalho Ininterrupto:** Agentes não são programas que “rodam” quando chamados pelo usuário. Agentes são sistemas que funcionam de forma ininterrupta e que percebem no ambiente as condições as quais foram concebidos a atuar.
- **Estado de Espera:** Agentes se mantem em um estado de espera (*idle*) percebendo os elementos do ambiente que ativem sua condição de ação;
- **Estado de Execução:** Agentes não são ativados para execução diretamente pelos usuários, e sim quando as condições de ação (gatilho) são observadas no ambiente;
- **Autonomia:** A interação com algum tipo de usuário não é necessária;
- **Iniciador de Processos:** O Agente pode necessitar que ele, ou outro agente, execute alguma tarefa específica para atingir o objetivo;
- **Reatividade:** O agente deve reagir ao ambiente, seja executando suas ações, seja modificando o seu comportamento;
- **Sociais:** O agente pode se relacionar com outras entidades, sejam humanas ou outros agentes.

Na Ciência da Informação Brasileira, o termo agente foi trazido por Tolosa e Bordignon (1999), os quais traziam para a área conceitos de agentes de *software* e o apresentaram como um objeto de pesquisa relevante para a CI, e como as discussões sobre este poderia prosseguir. Desde então não foram muitos os trabalhos que se debruçaram sobre os agentes de *software* na área, onde o uso do termo agente não se mostrava problemático.

Agentes de *Software* já estão presentes nos processos informacionais que ocorrem no mundo eletrônico. Então, mais do que um objeto de pesquisa, estes são protagonistas na relação sociedade-informação e são capazes de modificar hábitos (ou até mesmo culturas), o que ressalta a importância de se investigar as consequências das ações destes agentes em meios virtuais/eletrônicos.

É possível relacionar os agentes de *software* com alguns dos elementos previamente elencados. Facilmente se pode observar a semelhança entre os conceitos de agentes de *software* e os autômatos. Ambos são concebidos para se ter o máximo de autonomia, de forma ininterrupta e sem supervisão de um operador.

Alguns agentes podem possuir características sociais e assim se tornarem algo equivalente aos andróides simulacros. A partir do momento que um agente tenta simular o comportamento humano em um ambiente virtual, ele age como um andróide naquele espaço.

E por fim, agentes também podem colaborar com humanos de forma profunda o suficiente para compor um ecossistema cibernético, onde parte do problema é analisado e solucionado por seres humanos e outra parte por agentes. Então, essa composição “homem-máquina” sugere uma espécie cibernética baseado em comportamentos inteligentes e não em melhorias físicas.

Sumarizando, os agentes de *software* são entidades que residem no mundo virtual, ou seja, não é comum se caracterizar pelo uso de um *hardware* específico (a não ser o do próprio computador). Estes também são dotados de autonomia e percepção para atuar no contexto em que foram criados.

Ultimamente, outra classe de *software* vem ganhando a atenção da indústria, academia e meios de comunicação pela ativa participação na Internet e processos virtuais. Estes atores, chamados *bots*, tem chamado a atenção pela capacidade de interagir com pessoas na rede e prover serviços complexos na Internet. Mas serão os *bots*, *software* de uma categoria diferente dos agentes?

3.7 Bots

Esta categoria de *software* vem ganhando destaque, não só em contextos acadêmicos como também na imprensa. O termo foi cunhado, segundo o Dicionário Oxford¹², na década de 1960 e representa um diminutivo da palavra *robot*. Hodiernamente, esta expressão se refere somente aos programas que atuam na Internet e que são denominados como *Softbots* (*Software Robots*). Atualmente o termo *bot* é utilizado com muita frequência e este é corriqueiramente confundido com o vocábulo agente (GEMROT; CROM; PLCH, 2010).

Burkhardt (2018) afirma que *bots* são programas de *software* automáticos que executam tarefas repetitivas para coletar dados da Internet. Estes podem automatizar

¹² <https://en.oxforddictionaries.com/definition/bot>

processos repetitivos e demorados de forma eficiente, mas também podem ser implantados para minerar dados dos usuários ou para manipular a opinião pública. Dunham e Melnick (2008) afirmam que *bots* são aplicativos de *software* que executam tarefas automatizadas (*scripts*) pela Internet. Eles executam tarefas que são simples e estruturalmente repetitivas, em velocidade muito maior do que as pessoas.

Os *bots* vêm ganhando a atenção da comunidade científica pela sua significativa participação na Internet. A Empresa Globaldots¹³ aponta que 42% de todo o tráfego na “*surface web*” é realizado por algum tipo de robô, ou seja, quase a metade de todo o fluxo de informação na rede possui a participação de algum tipo de *bot*.

Gorwa e Guilbeault (2018) apresentam algumas classificações para os *bots*. Inicialmente eles os dividem em duas categorias: (i) *bots* bons e (ii) *bots* maliciosos, e assim são classificados de acordo com a finalidade de cada um. Os considerados benéficos são aqueles que trazem algum tipo de ganho para os envolvidos no contexto do sistema e normalmente costumam seguir algum tipo de legislação/norma. O outro grupo age realizando ações que são irregulares/ilegais ou que quebrem algum tipo de norma ética ou moral e, normalmente, estes não estão autorizados pelos usuários a agir.

Outra classificação trazida pelos autores é dada pelo modo de funcionamento do programa. Assim os autores reportam os seguintes tipos:

- **Crawlers e Scrappers:** São robôs que atuam na Internet executando um roteiro (script) pré-determinado e que normalmente não interagem com humanos. Podem ser bons ou maliciosos.
- **Chatbots:** São *bots* concebidos para se comunicar com seres humanos em interfaces de conversação, seja por texto ou por voz. São bons em quase sua totalidade.
- **Spambots:** São programas de divulgação de conteúdo que, normalmente, são indesejados pelo usuário. Estes programas, em geral, são maléficos por causar importunação aos usuários.
- **Socialbots:** São *bots* criados para interagir com os usuários em plataformas sociais e que, por vezes, tentam assumir uma identidade humana, não

¹³ <http://goo.gl/JGW5TA>

necessariamente para enganar o usuário, mas para se tornar mais amigável. Estes programas podem ser bons ou maléficos.

- **Ciborgues:** São estruturas híbridas de pessoas reais e sistemas automatizados que colaboram para realizar atividades na Internet. Esse ecossistema parte homem, parte máquina, remete ao termo ciborgue originalmente concebido. Estes sistemas podem ser bons ou maléficos.

Aparentemente, alguns autores utilizam a definição do termo *bot*, sem a evolução destes, se referindo apenas a alguns dos tipos (*crawlers*, *scrapers* e *spambots*) que têm o comportamento automatizado para fins específicos e não são tão elaborados.

Por outro lado, Gorwa e Guilbeault (2018) apresentam os *bots* dentro da categoria programas mais complexos, como *chatbots* e *socialbots*, que já não se enquadram na ideia de sistemas que executam *scripts* pré-definidos e de forma automática. Ao considerar esta nova visão, os *bots* se mostram bastante complexos, apresentando comportamentos inteligentes e capazes de socialização que não é comum para sistemas baseados em roteiros automatizados (*scripts*).

Gemrot, Crom e Plich (2010) apresentam algumas diferenças entre agentes e *bots*. Os autores afirmam que, embora possam existir interseções entre os conceitos e que comumente estes são confundidos e mal utilizados, tratam-se de entidades bem distintas. Segundo eles:

Bots nem sempre representam alguém: Se uma das características principais dos agentes é o reconhecimento de um responsável que responda pelas suas ações, e que o agente o representa, em um *bot* essa característica nem sempre é necessária. Quando se considera especificamente os *bots* maliciosos, estes não possuem uma entidade a quem representar. Deve-se compreender que “representar” é diferente de “obedecer”, pois é bastante claro que o objetivo de um agente é determinado por alguém, mas tal *software* não necessariamente precisa representar este mestre.

Bots precisam passar credibilidade: Muitos dos *bots* precisam ser capazes de passar uma “ilusão da realidade” para atingir os seus objetivos. Essa característica não pertence apenas aos *bots* que interagem diretamente com os humanos, tais como *chatbots* e *socialbots*. Existem *Spambots* capazes de escolher narrativas para interagir com alguém e há também os *crawlers* que modelam seu comportamento de busca e

recuperação da informação para agir como humanos e assim não serem reconhecidos. Os agentes não consideram estas restrições.

Bots podem precisar assumir uma identidade humana: Para aumentar a sua eficácia alguns *bots* podem precisar criar estratégias para serem confundidos com humanos reais. Essa característica é muito latente para *chatbots*, *socialbots* e *gamebots* (programas criados para atuar em jogos digitais), entretanto alguns *crawlers* já conseguem perceber sistemas de Captcha (“eu não sou um robô”) e tentam resolvê-los.

Além destas características listadas por Gemrot, Crom e Plch (2010), pode-se destacar algumas outras, dentre as quais:

Bots podem ser inicializados: Quanto aos agentes, atuam somente quando percebem determinadas condições no ambiente, um *bot* pode ser inicializado por um usuário humano ou *software*. Alguns *bots* podem realizar tarefas que pode ser disparadas (ou inicializada) por um outro ator, assim, *bots* podem tanto atuar a partir de condições (como agentes), bem como a partir de um comando externo.

Bots podem não perceber o ambiente: Alguns *bots* podem simplesmente agir sobre um determinado ambiente sem perceber as suas condições. Normalmente, *bots* que fazem tarefas repetitivas e automatizadas (baseados *scripts*) não precisam perceber ou conhecer característica do ecossistema onde estão atuando.

Bots podem estar inativos: Já foi identificado que *bots* podem ser inicializados por agentes externos, neste sentido, o *bot* pode não precisar estar ativo até realmente começar a execução. Os agentes estão ativos o tempo inteiro percebendo o ambiente, já os *bots*, nem todos precisam ter esta característica.

Assim, é possível considerar os *bots* como programas de computador que atuam no ambiente da Internet e interagem com os dados lá disponíveis para atingir objetivos de variados níveis de complexidade. Características distintas como autonomia, interação com usuários, colaboração, aprendizado e inteligência podem não ser observadas em todos os tipos de *bots*, embora, existem aqueles que apresentam todas estas especificidades, simultaneamente ou apenas uma delas.

Bots podem realizar desde atividades muito simples, como coletar informação em páginas da Internet (*crawlers*), como convencer um usuário humano a realizar uma

compra em uma loja virtual (*socialbot*) e é justamente por conta desta diversidade que fica difícil reunir características comuns a todos eles.

O que se pode observar é a evolução do termo, enquanto que no início o uso da palavra *bot* (ao se referir a classe de *softbots*) estava voltado aos sistemas de execução automática (*scripts*), – ou seja, de *software* que agem sem nenhuma autonomia e percepção de contexto – atualmente, se observa que os *bots* possuem características de inteligência e sociabilidade, e talvez por isso, a expressão tenha sido escolhida para representar tais tipos de *software*, que eram bastante complexos, mas que não tinham os mesmos objetivos de um agente, como observado por Gemrot, Crom e Plch (2010). Entretanto, acredita-se que o termo se popularizou e se firmou no seu sentido moderno.

Contudo, para os interesses desta pesquisa, se busca pela entidade de *software* que é capaz de se conectar a pessoas, compreender como se deve interagir com humanos, agindo de modo a influenciar opiniões e até mesmo moldar comportamentos. Ela certamente precisa ser composta de características presentes tanto nos *bots*, como nos agentes, mas que ao mesmo tempo, extrapolam os atributos individuais destes. Estes elementos, são encontrados nas Máquinas Sociais.

3.8 Máquinas Sociais

A ideia de Máquinas Sociais que permeia o senso comum está imersa em uma série de fenômenos que começaram a ser observados recentemente, razão pela qual é possível constatar uma ampla variedade de estudos que tendem a propor a sua própria definição do termo, o que impede de alcançar uma elucidação precisa e exata, mas, por outro lado, aponta diversas características que envolvem estes atores.

Tim Berners-Lee, o inventor da *World Wide Web*, e Fischetti cunharam o termo “Máquinas Sociais”, em meados de 1999, definindo-as como “uma construção sociotécnica onde as máquinas ajudam os seres humanos a aumentar a criatividade, facilitar a colaboração, e aumentar a conectividade global” (BERNERS-LEE; FISCHETTI, 1999 *apud* LEBOW, 2018, p. 172, tradução nossa). Na concepção de Bernes-Lee e Fischetti (2001), nota-se que as máquinas sociais são vistas como processos onde há colaboração das pessoas (humanos) responsáveis pela elaboração do trabalho criativo

do homem e a administração desses ambientes pelas máquinas, resultando em “novas formas de interação social e coordenação por meio de uma variedade de aplicativos baseados na Web” (LEBOW, 2018, p. 202, tradução nossa).

Berners-Lee; Fischetti (2001) ainda sugerem que as máquinas sociais resultam de um processo híbrido, por relacionar seres biológicos e tecnológicos no mesmo ambiente da Internet, onde “atividades de alto nível, que aconteceriam apenas dentro do cérebro de um humano, ocorrerão entre grupos ainda maiores e interconectados de pessoas agindo como se compartilhassem um cérebro intuitivo maior” (BERNERS-LEE; FISCHETT, 2001, p. 201-202, tradução nossa). Essa visão também é corroborada por Pierre Levy (2007), mas foi denominada por ele como inteligência coletiva. Então pode-se perceber que a gênese da ideia de Máquinas Sociais não é distante de uma visão discutida, amplamente, na Ciência da Informação.

As discussões sobre uma definição apropriada para máquinas sociais não se esgotaram nos autores que cunharam o termo, desde então verifica-se uma efervescência na academia para ampliar a aplicabilidade do termo “máquinas sociais”, principalmente quando a expressão tem sido usada com um significado diferente, representando máquinas responsáveis pela socialização da informação entre as comunidades, ou seja, uma intersecção das áreas e estudos de comportamento social e sistemas computacionais.

Roush (2005) sugere que a primeira aparição do termo na ciência é mais voltada ao lado social e menos a relação homem -> máquina. No texto intitulado “*Computing means Connecting*”, o autor descrevia uma máquina social como espaço de interação utilizado por um ser humano que é responsável pela socialização daquelas comunidades. Nesse toar, o autor posiciona as máquinas sociais como um produto da *web 3.0* que é considerada “uma plataforma para publicação pessoal e *software* social. Os exemplos são os sites informativos, como blogs, *craigslist* e Wikipedia” (ROUSH, 2005, p. 1, tradução nossa).

Dessa forma, as máquinas sociais são espaços colaborativos onde se formam infovias construídas por seus usuários em comunicação entre si ou com outras máquinas, a partir da troca de experiências, informações, percepções e outros dados sociais. Sendo assim,

Quanto mais pessoas usarem os novos serviços, mais poderosos esses serviços se tornam. Isso é porque eles são todos baseados em cooperação: as pessoas geralmente ficam felizes em compartilhar seus conhecimentos, experiências, criações, horários e locais, isso significa que eles podem aprender o que para as pessoas são importantes, estão pensando e fazendo (ROUSH, 2005, p. 1, tradução nossa).

O autor sugere que essa tecnologia seria uma espécie de “computação da nossa experiência cotidiana” (ROUSH, 2005, p. 1, tradução nossa) que resulta da exploração de padrões a partir da coletividade de pessoas que contribuem nesses novos espaços sociais, é como se a tecnologia estivesse sempre ativada e suavemente introduzida em nossas vidas, coletando nossos rastros e dando a esses “pedaços”, até então desorganizados, significados.

O que se percebe a partir da ideia de Roush é que o autor que popularizou o uso da expressão, e que as máquinas sociais não se confundem com as redes sociais virtuais, apesar da proximidade dos termos, conforme pontua Meira *et al.* (2011), visto que a primeira encontra-se no âmbito da *web 3.0* conhecida como a *Internet das máquinas sociais*, e a segunda reside no campo da *web 2.0*, também chamada de *web social*, quando surgem redes como o Facebook, Twitter e Fotolog (SANTANA *et al.*, 2013).

Nesse sentido, Santana *et al.* (2013, p.3) esclarece que “as máquinas sociais são plataformas programáveis em rede, cuja função e propósito podem ser, em boa parte, estendidos e redefinidos por quem detenha conhecimento para tal”, enquanto que as redes sociais virtuais são espaços onde “os usuários poderiam se conectar uns com os outros e a partir daí trocar informações” (SANTANA *et al.*, 2013, p. 2).

De um ponto de vista mais técnico, Meira *et al.* (2011) conceitua as máquinas sociais como entidades sociáveis que possuem unidades de processamentos onde são recebidas, processadas, produzidas e entregues informações (insumos), que se fazem através do contato com outras máquinas.

Já Burégio, Meira e Rosa (2013) aproximam a ideia de máquinas sociais à computação social, sendo aquelas a intersecção entre os *softwares* sociais, pessoas como unidades computacionais e *softwares* como entidades sociáveis. Nesse toar, considera Santana *et al.* (2013, p. 8) que as “máquinas e os usuários podem

simultaneamente ser membros de várias máquinas sociais diferentes e que cada máquina social é delimitada por sua finalidade”.

No livro *The theory and practice of online learning*, Terry Anderson resolve conceituar máquinas sociais, aproximando-se da visão trazida por Berners-Lee quanto ao seu caráter híbrido, revelando que esta é:

Uma comunidade de organismos vivos e híbridos cujos fracassos e sucessos informam as novas gerações de sucessores. Máquinas Sociais funcionam nos contextos de outras máquinas sociais (e redes sociais, redes tecnológicas e outros tipos de atores) (ANDERSON, 2008, p. 19, tradução nossa).

Nesse sentido, esclarece que uma máquina pode revelar seu caráter de sociabilidade onde observa-se seu potencial através da narrativa, reatividade, interatividade e colaboração entre atores (humanos ou não) (ANDERSON, 2008). Apesar disso, alerta o autor que “máquinas sociais não são um grupo homogêneo” (ANDERSON, 2008, p. 33, tradução nossa), isso porque é possível constatar diferentes maneiras de interação entre máquinas e humanos, de modo que novos arranjos sociais possam surgir a partir dessas relações.

De acordo com o autor, não há uma definição firme e precisa para o termo, sugerindo, portanto, que:

Dentro de uma máquina social, qualquer um substitui cada resposta humana no sistema, um por um, por um agente artificial, ou um substitui cada agente artificial, um por um, por um humano, e então faça a pergunta "quando isso deixou de ser uma máquina social híbrida e se tornou um sistema multiagente / interação puramente social?" (ANDERSON, 2008, p. 39, tradução nossa).

A visão de Anderson (2008) pode ser relacionada ao fato de que uma máquina social pode se comportar como ferramenta, processo, ambiente, apresentando múltiplas faces da tecnologia e inovação a depender do contexto em que é estudada ou utilizada.

Encarnação (2010), ao conceituar máquinas sociais, posiciona essas como aplicações, *softwares*, operações de infraestrutura, serviços, relacionamentos verificados na *web 3.0*, corroborando a ideia de que não haveria uma definição única do termo, mas sim uma diversidade de visões e atribuições ao mesmo objeto, assim “máquina social é uma aplicação *web* construída e projetada em rede que possui relacionamento com outras aplicações e serviços pré-existentes que a compõe, sendo

tais aplicações e serviços partes integrantes do sistema como um todo” (ENCARNAÇÃO, 2010, p. 38, tradução nossa).

Assim, ‘máquina’ está relacionada aos componentes operacionais do sistema (*hardware* e *software*), enquanto que ‘sociais’ refere-se a capacidade de interações e relacionamento que colaboram para a construção de uma máquina social que está organizada em rede. Desta forma, “máquinas sociais são entidades (aplicação, serviços, sistemas) disponibilizadas na rede que possuem relacionamentos entre si” (ENCARNAÇÃO, 2010, p. 38, tradução nossa), sendo assim, seriam entidades computacionais “projetadas para estarem em rede e disponíveis na forma das funcionalidades que elas provêm para outras máquinas” (ENCARNAÇÃO, 2010, p. 38, tradução nossa).

Shadbolt *et al.* (2016) apresenta outra visão de que as primeiras máquinas sociais eram programadas para coordenar a participação humana nesses ambientes, sendo que sua evolução resultou no surgimento de sistemas que são socialmente contextualizados e que servem como ferramentas para aquisição do conhecimento.

Desta forma, o próprio autor aponta para um universo composto por uma infinidade de máquinas com funcionalidades e propósitos diferentes, tais como o *Google* (modelo de negócio para mecanismo de buscas); *Galaxy Zoo* (*crowdsourcing* de classificação de imagens astronômicas); Sites de Mídias Sociais, tais como o *Facebook* e *Twitter* (plataformas que permitem comunicação e compartilhamento de informações entre indivíduos) e o Wikipédia (plataforma de co-criação para registro do conhecimento enciclopédico) (SHADBOLT, 2013).

É possível reconhecer que a computação social, ou seja, a computação que envolve dados sociais, é a base do desenvolvimento dessas máquinas, posto o sentido de colaboração e interação necessário para seu desenvolvimento. Desta forma, Smart e Shadbolt (2015) afirmam que este seria um sistema sociotécnico, onde elementos humanos e tecnológicos são participantes que colaboram entre si, construindo aquele universo. Completa Palermos (2017, p. 975) que as Máquinas Sociais se referem às “plataformas on-line que permitem que seus usuários interajam mutuamente uns com os outros”.

Outra característica deste sistema é a possibilidade dos participantes de desenvolver seu design e criação de uma forma *bottom-up* (usuário -> desenvolvedor) seguindo o contra fluxo dos sistemas tradicionais. Nesse contexto, “traços digitais da atividade humana se tornam cada vez mais disponível, adaptável e baseado na aprendizagem de máquina, sistemas podem usar esses traços para começar a coordenar a ação máquina-homem nas relações em rede” (SHADBOLT *et al.*, 2016, p. 108, tradução nossa).

Semmelhack (2013, p. 18) rememora que “todo o conceito de máquinas sociais é simplesmente uma extensão do contrato social que a maioria de nós já fizemos com um poderoso dispositivo eletrônico que transportamos todos os dias – nossos telefones celulares”. Para o autor, quando se fala de Máquinas Sociais, refere-se a um ambiente onde “incluem não apenas outras pessoas, mas dispositivos eletrônicos conectados à Internet, máquinas e engenhocas de todos os tipos” (SEMMElhACK, 2013, p. 16, tradução nossa), ou seja, trata-se de um ecossistema que envolve pessoas, organizações e máquinas colaboradoras entre si, que criam novos arranjos e formas de sociabilidade.

Para Semmelhack (2013), a interação não é apenas um elemento característico da máquina social, trata-se também de um fator que apontará para a sua capacidade de compartilhamento e conexão com o mundo para que os seus usuários se sintam cada vez mais próximos da realidade virtual e confiantes, razão pela qual cada vez mais ficamos “acostumados a ter um universo de informações em nosso bolso/ bolsa na forma de um telefone celular que, quando estamos sem ele, nos sentimos vulneráveis, despreparados ou até mesmo despidos” (SEMMElhACK, 2013, p. 67).

Outro fator que traz o sucesso e popularização das Máquinas Sociais para Semmelhack (2013) é o fato de conter em suas plataformas dados disponíveis e que sejam de interesse do seu público-alvo, contribuindo para que essas ferramentas se tornem conduto e catalisador de informações, aumentando, inclusive, os seus níveis de sociabilidade.

Roure *et al.* (2015) também entendem que o projeto sociotécnico que conta com a contribuição de elementos tecnológicos e agentes humanos é caracterizado como uma Máquina Social, sugerindo inclusive um nível de aperfeiçoamento das máquinas, de

modo que seja possível a comunicação entre elas, saltando da ideia trivial de que esses espaços são formados necessariamente pela interação humano e máquina. Assim:

Uma máquina não é apenas um computador que tem alguns usuários, mas sim é um projeto sociotécnico propositalmente composto por máquinas e pessoas. Podemos então ver o ecossistema como um conjunto de Máquinas Sociais em interação, em vez de camada de computadores utilizados por uma camada de seres humanos (Roure *et al.*, 2015, p. 2, tradução nossa).

Ao observar uma Máquina Social, deve-se atentar para o elemento da interação, responsável por distinguir a capacidade dessas máquinas daquelas tradicionais, posto que, nesse ambiente, verifica-se a comunicação mediada por humanos, computadores (automação) e máquinas sociais. Roure *et al.* (2015, p. 2) sugere que do diálogo entre máquinas surgem ecossistemas próprios, onde estão presentes forte influência exercida entre essas entidades e o aparecimento de relações desconhecidas pelos humanos.

O aspecto social é o produto desenvolvido por estas máquinas e é comum uma aproximação do mundo digital e físico, entretanto, apesar dessa semelhança apresentada pelas Máquinas Sociais, delimitá-las a um único sentido ou contexto resulta na exclusão de uma infinidade de outras máquinas, ou seja, é importante observar a constante mudança dos seus elementos integrantes. Nesse sentido,

Um dos aspectos fascinantes de uma Máquina Social como sistema é sua capacidade de resistir a tentativa de defini-la. Inspecione-o bem de perto e você acabará estudando detalhadamente suas partes constituintes (humanos, máquinas, *bots*); afaste-se demais e você perde de vista o que as partes constituintes estão fazendo. Uma das principais razões para isso é que uma máquina social de sucesso, intencionalmente projetada ou emergente, tem todas as propriedades de um sistema emergente; o tecido em si de suas partes constituintes sofre mutação sob sua influência mútua, assim como as interações entre eles (ROURE *et al.*, 2015, p. 3).

As Máquinas Sociais, para Santos (2015), são caracterizadas como *softwares* com motivações sociais, compostos por elementos humanos e computacionais, sendo “extensão da Web Semântica, criando o processo por meio do qual as pessoas executam as tarefas criativas e as máquinas realizam a administração dos dados” (SANTOS, 2015, p. 6).

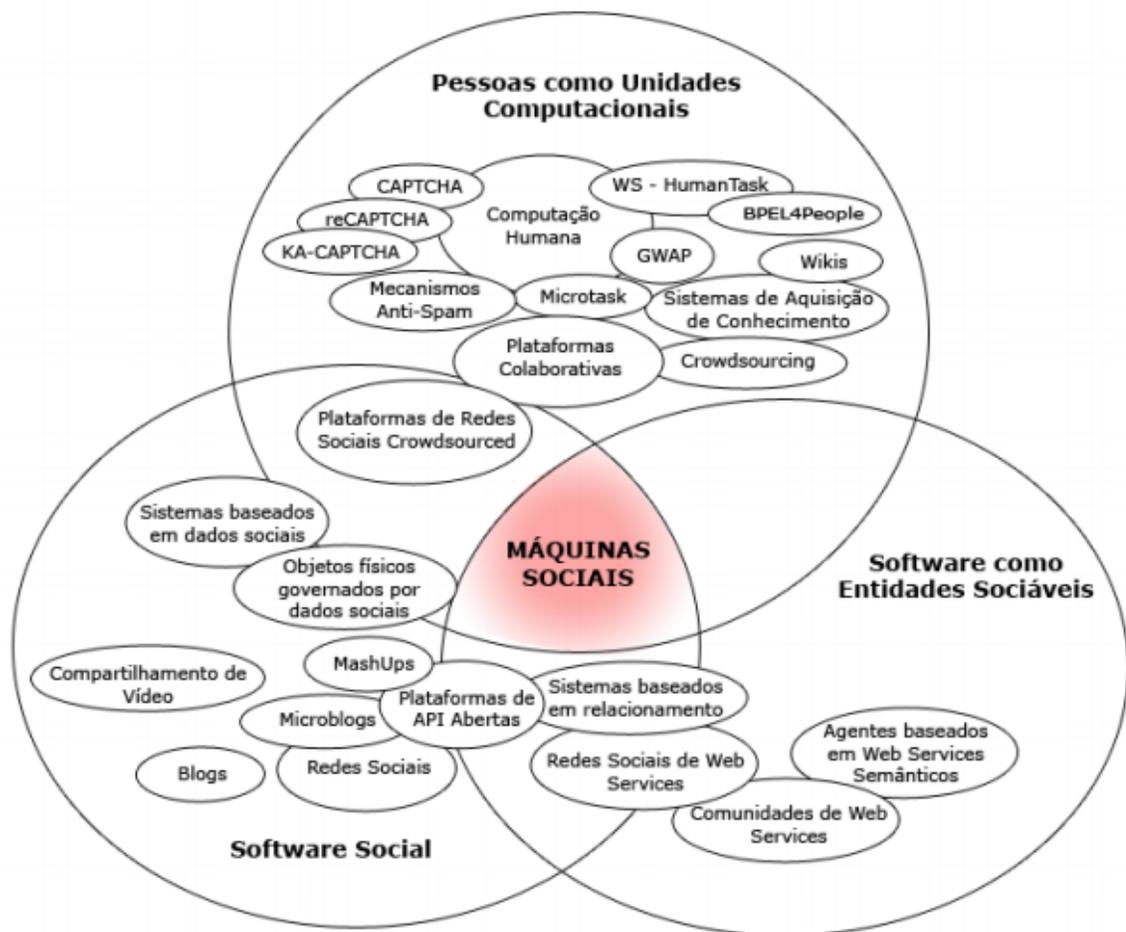
Ao conceituar Máquinas Sociais como “entidades sociáveis”, Santos (2015) aproxima-se da visão Meira *et al.* (2011) onde este define uma máquina social como uma entidade “conectável”. Ou seja, para Santos (2015), Máquinas Sociais podem ser

representadas como a intersecção de três pilares: **Pessoas Definidas como Unidades Computacionais; Software social e Associação de Software.**

Pessoas Definidas como unidades computacionais, a visão de pessoas envolve tarefas e comportamentos possíveis somente a seres humanos, tais como crowdsourcing; **Software social** Tecnologias para a Web que envolvem algum tipo de troca de informação entre usuários, tais como redes sociais da Web 2.0 e plataformas de API com Twitter e facebook; **Associação de software Softwares** e serviços para a Web que contêm uma ou mais camadas de integração com outros *softwares* sociais, tais como eles mesmos, outras aplicações de linked data ou redes sociais com outros webservices (SANTOS, 2015, p. 17).

O entendimento de Santos é baseado na concepção de Burégio, Meira e Rosa (2013), citado anteriormente, que observa a origem das Máquinas Sociais na computação social, sendo estas a evolução dos *softwares* que surgiram na *web 2.0*, baseadas, portanto em componentes formados pela intersecção desses três pilares, conforme exposto na Figura 1 a seguir.

Figura 1 - Máquinas Sociais e seus Componentes



Fonte: Santana; Lima (2018)

Pela análise da figura acima, é possível notar que a máquina social, na visão de Burégio, Meira e Rosa (2013) é um modelo que une máquinas e processos sociais, visando a interação entre indivíduos, grupos e máquinas com propósitos diversos que influenciam comportamentos, atitudes e resultados, no ambiente virtual e físico.

A ideia de uma máquina estática também é afastada na conceituação de Merchant (2016), que aponta para a construção contínua das Máquinas Sociais que conta com a participação ativa de usuários e máquinas, responsável pelo substrato informacional que resulta dessas relações. Assim define que “máquina social é um sistema tecnossocial complexo, composto de vários indivíduos ou grupos de indivíduos e componentes conectados através de uma plataforma em rede, modo predefinido de interação para um propósito particular” (MERCHANT, 2016, p. 787).

A capacidade de interação e o aspecto social que representa características de uma Máquina Social também são aderidas por Hendler e Mulvehill (2016, p. 10) que sinaliza para “à capacidade crescente que a inteligência artificial fornece para permitir que os computadores interajam de forma que tradicionalmente são consideradas apenas como o espaço social dos humanos”.

Meira (2011), ao se debruçar sobre o contexto das máquinas sociais, trata destas como entidades sociáveis, expandindo o sentido de sociabilidade para suas conexões com usuários humanos e outras máquinas:

Uma máquina social (MS) representa uma entidade conectável contendo um processamento interno unidade e uma interface de wrapper que aguarda solicitações de e responde (com respostas) a outras máquinas. Sua unidade de processamento recebe entradas, produz saídas e tem estados; e as suas conexões definem intermitente ou permanente relacionamentos com outras Máquinas Sociais, conexões que são estabelecidos sob conjuntos específicos de restrições (MEIRA, 2011, p. 4).

Considerando as diversas definições trazidas sobre o que seria Máquinas Sociais, é possível verificar que não há diversidade de conceitos sobre este objeto de estudo, porém é notável que muitos autores encontram alguns pilares base como interação, compartilhamento, informação em rede e conexão máquina- máquina / máquina-humano, sendo esses elementos comuns das SM (s).

Desta forma, para o presente trabalho, destaca-se as visões de Roure *et al.* (2015) e Burégio, Meira e Rosa (2013), por explicar melhor o tipo de *software* a ser investigado nesta pesquisa, tendo em vista observar as Máquinas Sociais como um conjunto formado de pessoas, *softwares* sociais e *softwares* sociáveis.

A visão de Roure *et al.* (2015) é interessante por este apontar que as máquinas sociais formam um ecossistema híbrido, onde os humanos dependem das máquinas e as máquinas dependem dos humanos para gerar um novo espaço informacional criado paulatinamente e de forma pouco disruptiva a ponto de a imersão humana neste espaço ser algo “quase natural”.

Já a visão de Burégio, Meira e Rosa (2013) auxilia a identificar os elementos que compõem uma máquina social (Figura 1) e posiciona diversos atores conhecidos na Ciência da Informação onde estes se encaixam em tal teoria. Ao definir as máquinas

sociais como a intersecção de três elementos, os autores deixam claro quais características, apresentadas de forma conjunta, uma máquina social deve possuir.

Segundo Recuero (2004, p. 7), o *Software Social* compreende “sistemas que visam proporcionar conexões entre as pessoas, gerando novos grupos e comunidades, simulando uma organização social”. Essa visão é complementada por Burégio, Meira e Rosa (2013) onde estes afirmam que, embora a ascensão destes tipos de sistemas tenha vindo com a *Web 2.0* (Blogs, Wikis e Youtube), a grande mudança trazida pelos *Software Sociais* é a capacidade em que estes têm de conectar coisas, não apenas sujeitos, e representá-las em forma de rede. Assim, nasce o conceito de entidades conectáveis que só são possíveis devido ao *Software Social*.

O *Software* como Entidade Sociável é definido por Nascimento *et al.* (2014) como aqueles que possuem a capacidade de tentar se conectar a outras entidades, a partir de seus elementos sociais, possibilitando a socialização deste com seres humanos e outros *softwares*. Assim, tem-se uma espécie de *software* que não só se reconhece em rede, mas reconhecem outros elementos em rede onde tentam colaborar.

Por fim, tem a Computação Humana que, segundo Goh e Lee (2011), sugere o caminho de mão dupla na relação entre homens e computadores, ao permitir que este último possa utilizar o auxílio do primeiro, para resolver problemas. Ou seja, pode existir a inversão do homem como operador da máquina, para um cenário onde a máquina pode utilizar a criatividade humana para solucionar problemas que o *software* não pode.

Desta forma, estamos diante de uma entidade de *software*, que se utiliza dos dados sociais disponíveis na rede e com uma complexa capacidade de se relacionar com outras entidades da rede (humanas ou não). É a partir desta definição que as máquinas sociais se encaixam no tipo específico de *software* capaz de fazer uma desinformação mais inteligente, capaz de ser apresentada aos seres humanos como algo feito por alguém semelhante, afastando a obviedade das mensagens automatizadas, classificadas como *spams* ou com traços de falsidade.

3.9 Comparação entre as Entidades

Foram apresentadas oito entidades que são classificadas, normalmente, como entidades não humanas, entretanto, existem nelas algumas características e peculiaridades que acabam por determinar o que o termo pretende destacar. Assim, se compreende que uma maneira de diferenciar tais entidades é observá-las segundo alguns parâmetros que são relevantes à, pelo menos, uma delas.

Para tanto, foram escolhidos oito atributos que se pode discorrer sobre cada um deles para cada entidade. Os atributos escolhidos foram:

- **Vontade Própria:** Capacidade de a entidade agir seguindo seu próprio livre arbítrio.
- **Autonomia:** Capacidade de a entidade agir sem a intervenção de outra entidade, realizando suas atividades sem supervisão.
- **Biológicos:** Se a entidade é um ser vivo ou artificial.
- **Controle:** Se refere à capacidade de configurar, ou parametrizar a entidade de modo que ela atue sobre padrões rígidos. Quanto maior o controle, maior será a possibilidade de configurar a máquina, bem como padronizar o funcionamento da máquina.
- **Cognição:** Se refere à capacidade que uma entidade tem em aprender, ou adaptar seu comportamento a um ambiente.
- **Hardware:** Se refere à necessidade de a entidade possuir algum corpo físico, tangível ou se esta pode ser uma entidade abstrata ou composta por elementos intangíveis (*software*, por exemplo).
- **Abstração:** Possibilidade de abstrair os detalhes de funcionamento de uma entidade, para simplificar a compreensão ou foco em outros aspectos.
- **Humanidade:** Capacidade que a entidade tem de se parecer com um humano, seja fisicamente ou em seu comportamento.
- **Sociabilidade:** Capacidade que a entidade tem em se sociabilizar e interagir com outras entidades de qualquer natureza.

Para cada uma dessas características foi atribuído um peso de 0 a 4 indicando o grau de importância daquele atributo para a sua definição. Desta forma, assim ficou a definição dos pesos para os atributos:

- **Vontade Própria:**
 - 0 (*Não Possui*) – A entidade não possui nenhuma vontade própria;

- 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
- 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
- 3 (*Alguma*) – A entidade pode possuir vontade própria em alguns casos;
- 4 (*Independente*) – A entidade não está sob controle de nenhum tipo.
- **Autonomia:**
 - 0 (*Operador On Line*) – A entidade precisa de um operador o tempo inteiro;
 - 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
 - 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
 - 3 (*Supervisão*) – A entidade pode precisar de intervenções esporádicas de um operador;
 - 4 (*Autônoma*) – A entidade não precisa de operadores.
- **Biológicos:**
 - 0 (*Artificial*) – A entidade não possui componentes biológicos;
 - 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
 - 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
 - 3 (*Híbridos*) – A entidade possui elementos biológicos e artificiais;
 - 4 (*Biológicos*) – A entidade é totalmente composta por elementos biológicos.
- **Controle:**
 - 0 (*Artificial*) – A entidade não pode ser configurada e padronizada;
 - 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
 - 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
 - 3 (*Parcialmente Configurável*) – A entidade possui muitos elementos configuráveis/parametrizáveis, mas ela não é 100% configurável;
 - 4 (*Totalmente configurável*) – Todas as ações e características de uma máquina são configuráveis, parametrizáveis e controlada pelo operador/programador.
- **Cognição:**
 - 0 (*Não Cognição*) – A entidade não apresenta características cognitivas (*inteligência, aprendizado, consciência e percepção*);
 - 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
 - 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
 - 3 (*Aprimoramento Parcial*) – A entidade desenvolve alguns elementos ligados a cognição de modo a não alterar o comportamento da mesma;
 - 4 (*Aprimoramento Total*) – A entidade desenvolve elementos cognitivos que não só melhoram sua capacidade de realizar a tarefa, mas também alteram seu comportamento.
- **Corpo Físico:**
 - 0 (*Nenhum*) – A entidade não precisa de um elemento tangível para existir;
 - 1 (*Irrelevante*) – Tal característica não é importante para o conceito;
 - 2 (*Opcional*) – A entidade pode ou não apresentar a característica;
 - 3 (*Tende a possuir*) – A entidade normalmente possui alguma espécie de corpo físico;
 - 4 (*Sempre*) – A entidade sempre apresenta um corpo físico.
- **Abstração:**

- *0 (Especialista) – O operador deve ser especialista no funcionamento da entidade para operá-la;*
- *1 (Irrelevante) – Tal característica não é importante para o conceito;*
- *2 (Opcional) – A entidade pode ou não apresenta a característica;*
- *3 (Interface) – O operador não deve ser especialista no funcionamento da entidade para operá-la, mas existem algum tipo de interface para utilizá-la;*
- *4 (Abstrata) – A entidade é totalmente operável por pessoas que não conheçam o funcionamento da máquina e o uso de interfaces é mínimo.*
- **Humanidade:**
 - *0 (Diferente) – A entidade não se parece nem foi concebida para se parecer com seres humanos;*
 - *1 (Irrelevante) – Tal característica não é importante para o conceito;*
 - *2 (Opcional) – A entidade pode ou não apresentar a característica;*
 - *3 (Parecido) – A entidade é perceptivelmente diferente de um ser humano, mas algumas de suas características foram criadas com o intuito de se aproximar dos seres humanos;*
 - *4 (Idêntica) – A entidade se parece com seres humanos a ponto de confundir outro humano, mesmo que em um primeiro contato.*
- **Sociabilidade:**
 - *0 (Isolada) – A entidade funciona de maneira isolada e não se comunica com outras entidades;*
 - *1 (Irrelevante) – Tal característica não é importante para o conceito;*
 - *2 (Opcional) – A entidade pode ou não apresentar a característica;*
 - *3 (Tende a ser) – A entidade pode não ser sociável, mas a grande maioria apresenta esse comportamento;*
 - *4 (Sempre é) – A entidade sempre se comunica/interage com outras entidades.*

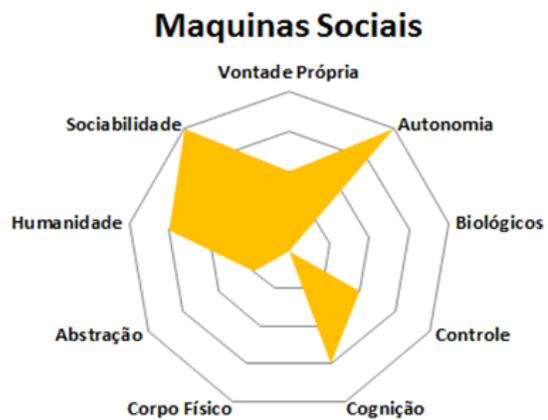
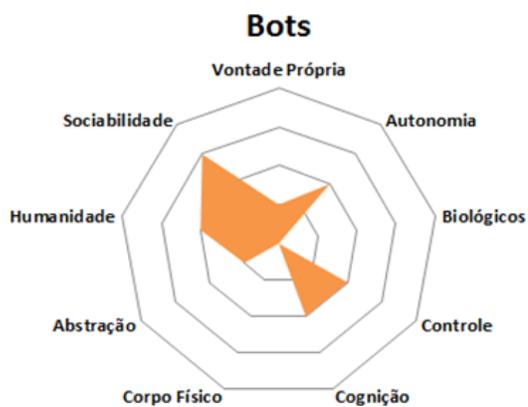
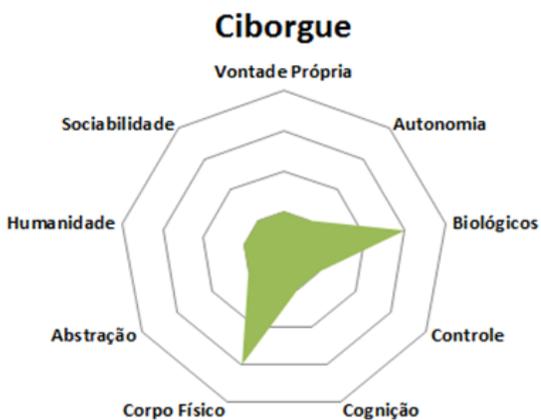
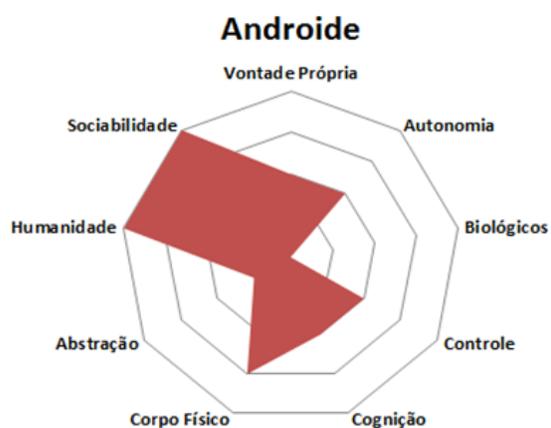
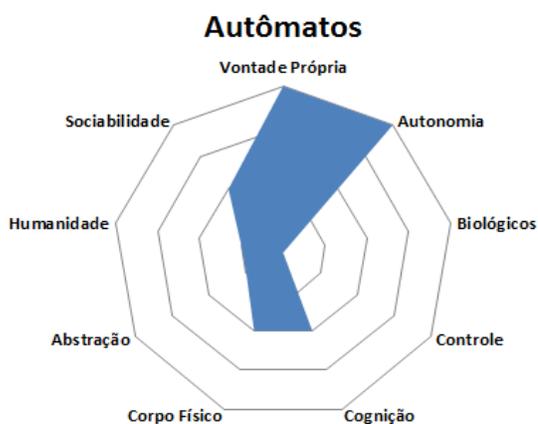
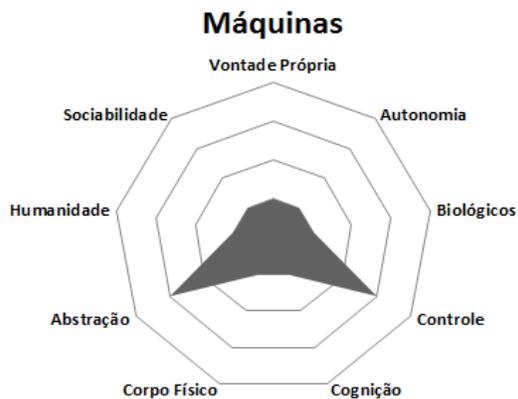
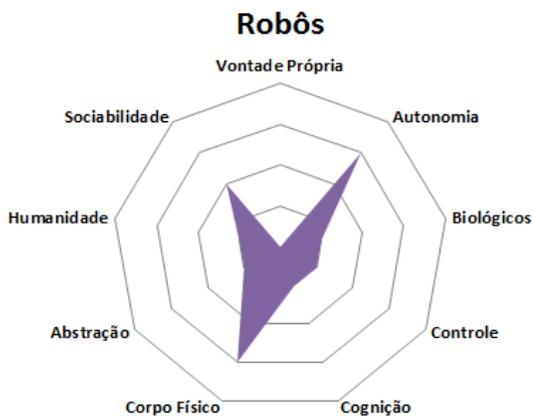
Para destacar as diferenças entre as entidades, foi atribuída, para cada uma delas, a classificação (de 0 a 4) para os oito atributos e gerado um gráfico de radar, o que permite fazer uma comparação visual, e rápida das mesmas. O resultado pode ser visto na Figura 2 a seguir.

Embora os termos sejam constantemente confundidos, pode-se perceber que eles são bem distintos e que estes, apesar de possuírem características comuns, não compartilham de um conjunto significativo delas de forma equânime, para se observar alguma semelhança.

Iniciando pelos **Robôs**, pode-se observar que a utilização geral do termo é para se referir a entidades não biológicas, que possuem um corpo mecânico capaz de realizar algum tipo de movimento e que executam as tarefas programadas por seres humanos.

Esse senso comum do que seria um robô difere da sua definição original que é a criação de entidades que realizam trabalhos forçados.

Figura 2 - Comparação entre as entidades



Fonte: Autoria Própria

A primeira divergência é que, na obra original, os robôs são entidades biológicas e não eletromecânicas, como se costuma visualizá-los. Isso pressupõe a existência de um corpo (*hardware*) que os materializa, mas se torna mais fácil para os seres humanos pensar que o grau de subserviência esperado de um robô, seja atribuído a um ente sem vida, ou pelo menos, sem autoconsciência, e talvez por isso seja remetido a figuras mecânicas para caracterizar robôs.

A segunda divergência é que, na obra original, os robôs não seguem uma programação estrita para suas ações, eles são orientados a objetivos e assim podem atingir, teoricamente, uma mesma meta de formas diferentes. Entretanto, esta diferença apenas reforça que tais entes devem ser autônomos, funcionando com o mínimo de intervenção humana.

Então estas são as duas principais características ligadas aos robôs: **Corpo Físico e Autonomia**. Robôs são entidades que existem no mundo real e trabalham com o mínimo, ou as vezes sem nenhuma, supervisão humana para realizar as suas tarefas. A estas duas características foram atribuídas nota 3 (em um máximo de 4), pois tanto existem robôs que ainda precisam de pouca supervisão humana (*Human on the loop – HOTL*) quanto definições de robôs feitos de *software* (*softbots* ou *bots*), afastando assim a nota máxima para estas características.

Robôs podem, ou não, exercer a **sociabilidade** (Nota 2), entretanto existe a compreensão que robôs precisam, de alguma maneira, ser informados dos seus objetivos e também, de alguma forma, efetuar suas ações no contexto em que ele existe. Então, é esperado algum grau de interação do robô com o mundo ao seu redor, mas não necessariamente haverá nesta relação a sociabilidade.

Os atributos **humanidade, biológicos, controle, cognição e abstração** são considerados como não relevantes para a definição de robôs (Nota 1). Se eles são parecidos ou não com humanos (humanidade), se possuem elementos biológicos ou mecânicos (biológicos), se estes podem ter seus comportamentos padronizados (controle), se estes apresentam características cognitivas (cognição) ou se estes representam apenas uma simplificação da forma de funcionamento (abstração) não são

importantes na concepção dos robôs. Vale ressaltar que um robô específico possui alguma faixa de graduação à todas estas características, mas elas são irrelevantes.

Por fim, em uma última característica, **vontade própria**, a nota seria 0, pois os robôs não possuem nenhum senso de autoconsciência ou mesmo de recompensa que não seja atingir seus objetivos definidos por alguém.

Prosseguindo para as **Máquinas**, observa-se que o termo é utilizado geralmente quando há intuito de reduzir um “mecanismo” de funcionamento complexo à sua funcionalidade. O senso comum é de que a máquina produz seus resultados/efeitos a partir de uma estrutura caixa preta, com um grau, normalmente elevado, de padronização. Esta visão para máquinas não é muito distante daquela que foi construída historicamente a partir da sua gênese oriunda do latim.

Existe uma visão, muito particular, de que as máquinas tratam-se de objetos eletromecânicos, cujo funcionamento interno se mostra bastante complexo, mas são manipulados por seres humanos a partir de interfaces simples que configuram a máquina para trabalhar de acordo com os parâmetros desejados. E, é esta visão que se pretende aqui extrapolar e mostrar que o conceito máquina abrange um escopo maior.

Assim, foram elencadas como características principais das máquinas a **abstração** e o **controle**. Máquinas sempre apresentam um grau de abstração sobre o seu funcionamento, de modo que o usuário/observador perceba com clareza o resultado da utilização ao invés de se perder nos detalhes do “como” tal resultado foi atingido. Máquinas abstraem sua complexidade em graus diferentes, sendo algumas muito simples de se utilizar, outras ainda continuam complexas e por isto foi considerada a nota 3 (de um máximo de 4). O mesmo se pode afirmar do controle, normalmente, nem todas as máquinas têm todos os seus parâmetros configuráveis e por isso também se atribui a nota 3.

Os outros atributos **sociabilidade**, **humanidade**, **biológicos**, **cognição**, **autonomia**, **corpo físico** e **vontade própria** são considerados como não relevantes para a definição de máquinas (Nota 1). Vale destacar o quesito autonomia, embora se tenha a tendência de visualizar máquinas cada vez mais autônomas, importa lembrar que os automóveis populares atuais são máquinas que precisam de supervisão

constante dos motoristas, encaixando-se na categoria *Human in the loop* – HITL, onde o controlador humano intervém em cada decisão.

Os **autômatos** já apresentam um uso mais restrito do termo, sendo utilizado em sua forma mais regular como qualquer objeto, máquina ou entidade não humana que possui alguma ação automatizada realizada como se fosse feita por um ser vivo. Tal tarefa pode ser algo simples, abrir portas e servir vinho, até mesmo se configurar em uma entidade completamente autônoma. Entretanto, vê-se que um autômato se refere a uma entidade, viva ou artificial, que possui vontade própria e autonomia suficiente para realizar suas ações sem supervisão humana.

Então, caso se queira tratar de criaturas artificiais que possuem vida própria, é deste tipo de entidade que se refere. Neste sentido, os autômatos apresentam o maior grau, nota 4, em **autonomia** e **vontade própria**. Nesta visão, pode-se dizer que os autômatos são os seres mais opostos aos robôs, indo de encontro as ideias apresentadas por Choi (2011) de que o autômato é uma evolução dos robôs e reforça o pensamento de Santana e Oliveira (2018) que utilizaram a metáfora do Mágico do Oz para questionar: o que o homem de lata deveria pedir para deixar de ser um robô? Um coração ou um cérebro? Ou, quem seria o robô, o homem de lata ou o espantalho?

Se os autômatos apresentam este grau de independência, não seria absurdo pensar que estes também apresentariam, um grau de **sociabilidade**, **cognição** e a necessidade de um **corpo físico** tangível. Assim como nos robôs, essas características nem sempre são obrigatórias em um autômato, ele pode ser recluso e sem características cognitivas relevantes como (Golem de Praga) ou mesmo não possuir corpo como aqueles presentes na teoria dos autômatos matemáticos. A partir desta consideração a estes atributos foi atribuída a nota 2.

Características como **humanidade**, **biológicos** e **abstração** são considerados indiferentes para estes autômatos. Ressalta-se, novamente, que os autômatos possuem seu grau de semelhança com seres humanos (humanidade), pode ser biológico ou não (biológicos) e podem esconder detalhes do funcionamento do ente (abstração), mas não são estas as características que determinam ou não ser uma criatura ou ser um autômato (Nota 1).

Por fim, destaca-se que um autômato não pode ser configurado/padronizado (**controle**) de forma tão estrita. Não se conhece histórias de autômatos (seres com vontade própria) tendo seus parâmetros definidos por entidades externas. Portanto, este é o único atributo com a nota 0 para esta entidade.

Seguindo para os **androides**, estes se mostram como entidades que tem como objetivo ser o mais parecido possível com um ser humano. Se no passado a intenção era recriar entidades, fisicamente, o mais parecido possível com os seres humanos, hoje se tenta replicar também os comportamentos. Por isso, estas entidades apresentam o grau máximo (Nota 4) para **sociabilidade** e **humanidade**. Assim, é possível observar uma convergência de características físicas e de relacionamento com outros seres humanos com o intuito de se parecer com um humano.

Vale ressaltar que os androides quase sempre apresentam um **corpo físico**, entretanto, existem novas formas de androides que tentam simular comportamentos humanos na Internet, embora estes não tenham corpo. Esta é uma característica atribuída a um *bot* (visto mais adiante), entretanto, estes são considerados androides por natureza. Assim, esta característica recebe a nota 3.

Já a **cognição**, **controle**, **autonomia** e **vontade própria** são elementos que podem não estar presentes nestas entidades. Por se tratar de entidades criadas para se passar por seres humanos, imagina-se que estes elementos são sempre obrigatórios em tais entidades, mas, podem existir androides que não as apresente. Neste sentido, tais entidades recebem a nota 2 para estas características. Já a **abstração** é um elemento que não define se um ente é ou não um androide, e por isso este é considerado aqui como indiferente (Nota 1).

Por fim, compreende-se que um androide sempre será uma estrutura não **biológica**. O que existe de tecnologia atual, não remete aos androides nenhuma característica biológica ou mesmo híbrida (ciborgues) construída para se parecer com seres humanos.

Já os **ciborgues** se mostram como entidades que tem como ênfase a hibridização de organismos biológicos e partes eletromecânicas com o intuito de aperfeiçoar as habilidades da entidade aproveitando as vantagens das estruturas biológicas “otimizadas” por peças artificiais. Para este conceito não se percebe a multiplicidade de

contextos, o que torna mais uniforme a compreensão do que se trata quando se vê a palavra ciborgue.

Por isso, para esta entidade foram consideradas as maiores notas (Nota 3) para as características **biológicos** e **corpo físico**. No ponto biológicos, o grau 3 representa justamente os seres híbridos, enquanto que, no corpo físico, este grau representa que tal entidade tende a possuir um corpo, entretanto, alguns sistemas de informação híbridos (humanos e máquinas) começam a ser chamados de ciborgues pela comunidade científica, trazendo este caráter intangível a esta entidade.

Todas as outras características: **cognição**, **controle**, **abstração**, **sociabilidade**, **humanidade**, **autonomia** e **vontade própria**, são irrelevantes para determinar se um ente é um ciborgue ou não. Por isto estes recebem o grau 1 como atribuição.

Seguindo para os **agentes de software**, estes são as primeiras entidades totalmente intangíveis apresentadas até aqui. Como principal característica, tais agentes independem de um operador humano para funcionamento e, por isso, ela tem como atributo máximo (nota 4) a característica **autonomia**. Por outro lado, existe uma série de características que este não possui, tais como: **cognição**, **controle**, **biológicos**, e **corpo físico**. Este não possui uma espécie de *hardware*, seja biológico ou eletrônico e apenas alguns de seus subtipos podem possuir algum grau de inteligência (cognição) e podem ter seus parâmetros ajustados para uso (controle).

Tais agentes podem ou não ser sociáveis (**sociabilidade**), e por isso foi atribuída a nota 2 para esta característica, enquanto, que são indiferentes as questões de **abstração** para tal entidade (nota 1).

Já os **bots** apresentam uma semelhança com o termo robô quando considerada a evolução do seu significado ao longo do tempo. Se na origem do termo, este era utilizado para representar *softwares* que executavam tarefas repetitivas e pré-programadas (*scripts*), os *bots* passaram a ser relacionados com alguns determinados tipos de agentes de *software*, tais como os *chatbots* e *socialbots*. Por se tratar de uma ampla gama de programas, hoje mais voltada a interação com humanos, pode-se compreender que os *bots* têm uma distinção dos agentes de *software*.

Inicialmente, eles têm uma participação cada vez maior no atendimento de usuários reais por sistemas automatizados. Por isso, o maior grau de **sociabilidade**

(Nota 3) é atribuído a tais entidades. Vale ressaltar que *scripts* de tratamento de dados, que não interage com outras entidades, também são considerados *bots* e por isso este atributo não tem a nota máxima.

Essa visão contemporânea dos *bots* como os interagentes em sistemas de informação também o remete a uma certa **autonomia** e **cognição**. Já que ele precisa “dialogar” com um usuário sem a supervisão de um ser humano e aprender sobre as preferências de tais clientes humanos, estas características passam a estar presentes em muitos destes *bots*. Eles também passam por um processo de se parecer mais com as pessoas (**humanidade**), tanto que as grandes empresas atribuem nomes aos seus *chatbots* como a Joice (Oi Telecom) e a Bia (Bradesco). Às vezes, alguns comportamentos devem ser também de seres humanos e tais *bots* podem demonstrar empatia, alegria e outras ações que remetam ao usuário (humano) a sensação de estar sendo atendido por um igual. E por fim, muitos destes *bots*, podem ser estritamente controlados para realizar as suas ações de modo mais eficiente e eficaz (**controle**). Normalmente, estes *bots* são os *scripts* realizando tarefas repetitivas e de objetivos bem específicos, sendo diferentes daqueles que lidam com pessoas. Para todas estas características, foi atribuída a nota 2 (podem ou não apresentar), por serem importantes para definir que tipo de *bot* está sendo apresentado.

Bots não são normalmente referenciados para esconder algum tipo de entidade complexa (**abstração**) e nem se mostra preponderante, dentro da classificação dos *bots*, se estes possuem algum tipo de **vontade própria**. Para estes dois itens a nota atribuída foi 1. Compreende-se que certamente tais entidades não são **biológicas** e nem possuem algum tipo de **corpo físico**, o que as qualificam como nota 0 para os *bots*.

As máquinas sociais, que são o objeto de estudo deste trabalho, podem ser classificadas como um subtipo de *bot*. Entretanto, o que se destaca é a capacidade que estas têm de se socializar (**sociabilidade**), seja com humanos seja com outros *bots* e agentes de *software*, e pela total independência de um operador humano (*human out of the loop*) em seu funcionamento (**autonomia**). Para estas duas características foram atribuídas o grau máximo (Nota 4).

Máquinas Sociais se destacam também pela sua semelhança com seres humanos (**humanidade**) e capacidade de ajustar seu comportamento (**cognição**) para atingir seus

objetivos de forma mais eficaz. Neste sentido, tais entidades tentam se aproximar ao máximo do comportamento humano para que não se crie um distanciamento dos usuários reais e, sempre que estas máquinas são atualizadas, elas revelam um comportamento diferente para melhor realizar suas atividades. Portanto, estes atributos ficaram com a Nota 3.

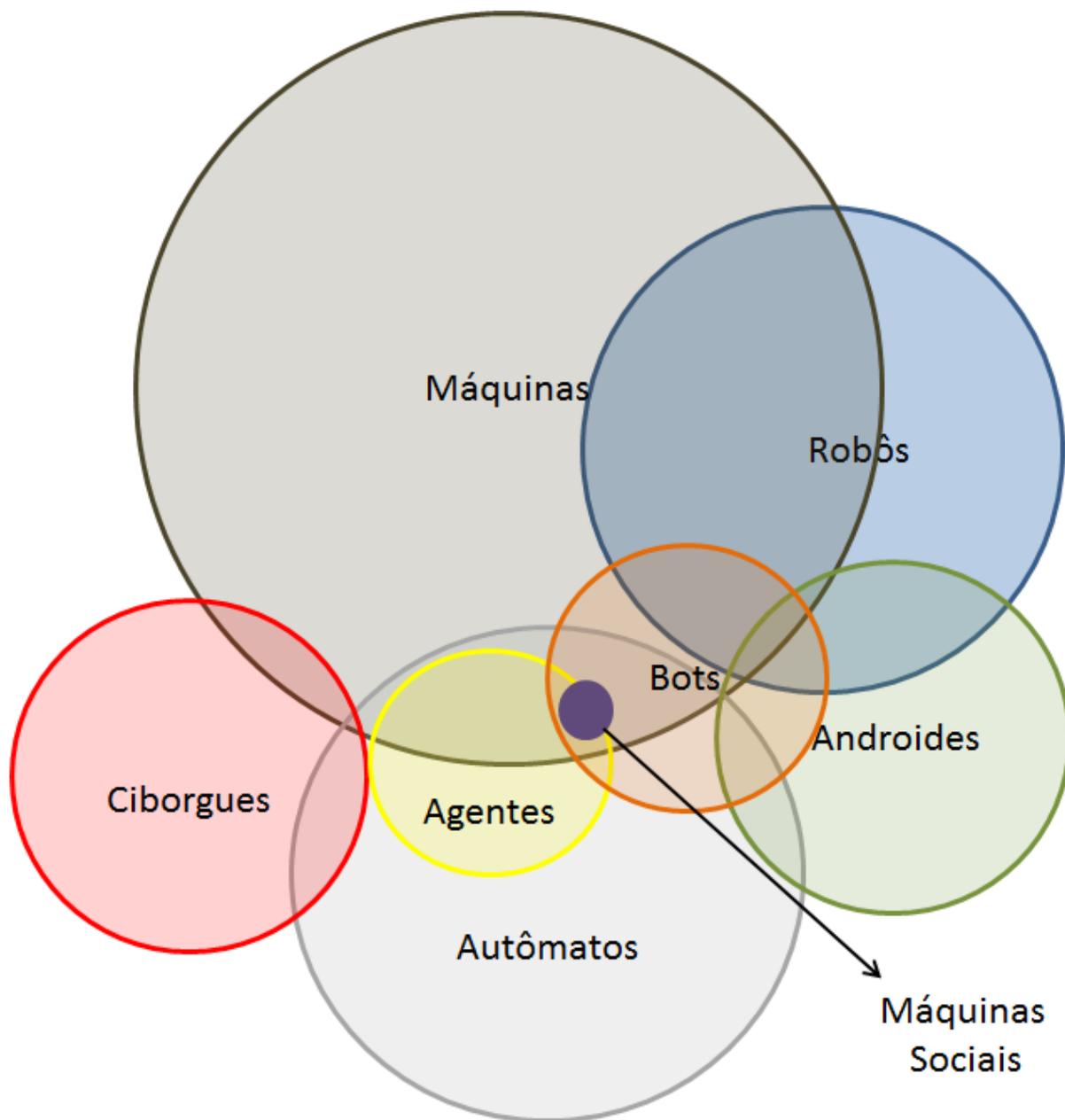
Tais máquinas podem possuir **vontade própria**, ou seja, ela própria pode definir suas metas bem como estas podem ser configuradas e ajustadas para trabalhar sob configurações pre-definidas pelos seus “programadores”. Estas características são opcionais e, portanto, se apresentam com a nota 2.

Assim como os *bots*, as máquinas sociais não são mencionadas como uma entidade complexa cujo principal objetivo é esconder uma complexidade (**abstração**), apesar do nome máquina, para este atributo foi atribuída a nota 1. E, também igualmente aos *bots*, tais entidades também não são **biológicas** e nem possuem algum tipo de **corpo físico**, o que as qualificam como nota 0.

As duas entidades mais semelhantes, visualmente falando, são os *bots* e as máquinas sociais, uma vez que as máquinas sociais são um tipo de *bot*, entretanto, mais especializado.

Por fim, a Figura 3, também em uma abordagem visual, apresenta a relação entre os objetos considerando suas semelhanças e diferenças.

Figura 3 - Relação entre as entidades



Fonte: Autoria Própria

As intersecções significam que existem objetos concretos que pertencem as duas classes simultaneamente. Por exemplo, a Intersecção entre Robôs e Máquinas, significa que existem máquinas que também são robôs, embora nem todas as máquinas sejam robôs e nem todos os robôs sejam vistos com o olhar de máquina, e a maioria dos robôs seja vista como máquinas. Entretanto, cabe refletir sobre as relações das máquinas sociais, uma vez que elas estão totalmente imersas nas categorias máquinas, autômatos e *bots*, e quase em sua totalidade está se situa na categoria agentes.

A visão de que uma máquina social é um ente abstrato (máquina) não é absoluta, mas não é papel da Ciência da Informação abrir a caixa preta de tais entidades e mostrar, com detalhes, o seu funcionamento. Desta forma, o olhar da CI para as máquinas sociais, deve ser a sua finalidade, impacto social, determinar sua participação em determinados regimes de informação e identificar as práticas informacionais em que estas estão presentes, abstraindo, o máximo possível o seu funcionamento.

Uma máquina social também está totalmente imersa na categoria autômato devido ao grau de autonomia que esta possui para atingir os objetivos determinados. Por mais que haja uma intenção, finalidade ou objetivo na execução de suas ações, esta “pré-programação” se dá em um nível muito alto de abstração, onde as táticas e estratégias adotadas pelo *software* em questão podem ser modificadas, redefinidas e até mesmo transformadas pela entidade. Então, neste contexto, embora haja uma pré-definição de um objetivo por parte do *software*, ele pode ser equivalente a índole ou moral de um ser humano. Ou seja, todo ser humano age de acordo com suas convicções ou objetivos de curto prazo, e todas as ações destes são voltadas para este objetivo maior. Então, hoje, o comportamento das máquinas sociais para seguir esta “bússola moral” não está tão distinto daquele apresentado pelos seres humanos em atingir seus objetivos.

Uma máquina social também se encaixa na categoria *bot* por se colocar como uma especialização deste. Uma máquina social é um “*socialbot*”, que trabalha a partir de dados sociais (em rede), possui algum nível de cognição (artificial) e que tem como objetivo interagir / socializar-se com outros membros da rede, sejam estes humanos ou não.

Por fim, existem muitas máquinas sociais que se comportam como agentes, embora não todas estas o sejam. Em geral as máquinas sociais possuem traços de agentes inteligentes (trabalham ininterruptamente, reagem ao ambiente, possuem características sociais), entretanto, nem todas são assim. Existem máquinas sociais que podem ser inicializadas (são executadas a partir de um comando humano), isso significa que eles podem estar inativos. Algumas máquinas sociais são criadas para simular uma pessoa com um intuito de engodo (e não de representação), outras não precisam perceber o ambiente ao redor e sim utilizar um conjunto de dados sociais (que podem

ser estáticos) para executar. Estas últimas são características mais próximas aos *bots* do que os agentes em si.

3.10 Síntese da Sessão

Esta seção teve como principal função diferenciar diversos elementos não biológicos que comumente são utilizados de forma indistinta pela academia e essa realidade não é exceção na Ciência da Informação. Apresentar cada uma das oito entidades, suas finalidades, semelhanças e diferenças ajudam a compreender o olhar que pode ser dado a cada uma delas e desmistificar certas atribuições que podem ser feitas a cada entidade, por exemplo: (i) seres humanos podem ser robôs; (ii) o R2/D2 da franquia guerra nas estrelas possui mais características dos autômatos do que de robôs e andróides; e (iii) agentes de *software* e *bots* não são, em muitos casos, a mesma entidade. Assim, esse posicionamento faz necessário destacar que a entidade que mais participa de fato do processo de desinformação são as máquinas sociais.

Outra função desta seção é apresentar o conceito de máquinas sociais e suas visões, e aproximá-lo de uma visão mais adequada a Ciência da Informação, ao mesmo tempo em que demonstra a diferenciação dos *bots* e agentes de *software*. Vez que, uma parte considerável dos *bots* e agentes não são eficientes para desinformar, enquanto, que as máquinas sociais, a partir de suas características, são a entidade mais “preparadas” para tal, inclusive para participar de complexos processos de análise e interação com usuários.

Por fim, esta seção tem como uma última finalidade propor que muitas das atribuições as entidades não biológicas discorrem muito mais de uma compreensão, ou falta dela, dos pesquisadores sobre tais entidades. Assim, é comum encontrar certas menções a robôs autônomos (autômatos), agentes de *software* que ficam inativos (*bots*) ou até mesmo de máquinas que possuem especificação detalhada (sem abstração). Estas observações, mais do que o preciosismo resulta no emprego equivocado do termo, revelam uma falta de cuidado sobre que aspecto do dispositivo (seja *hardware* ou *software*) que se quer discutir, e ao pôr em destaque as máquinas sociais se deixa claro para esta pesquisa que a entidade a ser observada é aquela que:

- É puramente *software*;
- Está em rede com outras entidades;
- É capaz de utilizar as informações disponíveis na rede (dados sociais) para atingir seus objetivos de forma mais eficaz;
- Apresenta um elevado grau de autonomia, e possivelmente, inteligência, para escolher / reprogramar suas táticas e estratégias para atingir tais objetivos;
- É capaz de se socializar, seja em um chat, seja em uma curadoria de conteúdo, ou mesmo em uma interação de mais longo prazo com outras entidades.

4 PROCEDIMENTOS METODOLÓGICOS

Para esta pesquisa foi utilizado um conjunto de métodos que trazem evidências complementares acerca do fenômeno estudado. A variedade de manifestações observadas no mundo real propicia a criação de um quadro mais completo que envolve a participação de atores não humanos, em especial as máquinas sociais, no processo de desinformação. Neste sentido, esta pesquisa apresenta, em etapas distintas, as **abordagens quantitativa e qualitativa** (MICHEL, 2009).

Em um primeiro plano, a abordagem quantitativa terá os seus dados originais retirados de publicações acadêmicas, jornalísticas e de sítios na Internet que tenham o foco em tecnologia. Embora os dados desta pesquisa sejam derivados de outras fontes (secundárias), é a partir deles que algumas das bases desta discussão serão fundamentadas. Segundo Michel (2009), a pesquisa quantitativa considera que “tudo pode ser quantificável”, isto significa que os dados apresentados na pesquisa fazem um maior sentido se visualizados em forma de números. Nesta abordagem, a coleta e análise dos dados possuem métodos mais apropriados para se trabalhar com quantidades e métodos mais ligados a matemática e estatística.

A aplicação de tais métodos é mais adequada quando se necessita de precisão nos resultados para que se minimizem os desvios e vieses, a partir do uso de indicadores preestabelecidos. Michel (2009, p. 6) afirma que na pesquisa quantitativa “os resultados são obtidos e comprovados pelo número de vezes e exatidão em que o fenômeno ocorre”. Ainda segundo a autora, neste tipo de pesquisa o resultado deve ser “numérico, exato e inquestionável”.

Assim, no presente trabalho, optou-se por utilizar-se dos diversos gráficos que apresentam (i) quantidades brutas de interações; (ii) relações de crescimento temporais; e (iii) outras correlações numéricas que auxiliam a caracterizar a participação de máquinas, bem como observar a reação (virtual) dos usuários humanos neste contexto.

Esta pesquisa também se caracteriza como qualitativa no sentido trazido por Michel (2009, p.12) por existir uma “relação dinâmica, particular e temporal entre o pesquisador e o objeto de estudo”. Pois, se faz necessária a interpretação dos fatos dentro de um contexto já que o “ambiente de vida real é a fonte direta para a obtenção

de dados”, tendo o entendimento do pesquisador, com o máximo de imparcialidade, como chave para significar as respostas.

Segundo a mesma autora, as pesquisas qualitativas são fundamentadas na discussão e correlação de dados interpessoais, e na coparticipação das situações dos informantes analisados a partir da significação que estes dão aos seus atos. Neste tipo de pesquisa a análise detalhada do fenômeno e a interpretação das evidências vão além da frieza das quantificações e da descontextualização do ambiente. O caráter qualitativo da pesquisa se dá pela interpretação das reações/comportamentos humanos diante das ações iniciadas por máquinas.

Quanto aos meios esta pesquisa também apresenta uma diversidade de facetas. Iniciando pela **pesquisa bibliográfica** que, segundo Gil (2008, p. 45), “auxilia no aprimoramento das ideias, descobertas de intuições e construções de hipóteses a respeito de um determinado problema”. Neste trabalho, conforme a classificação de Michel (2009), foi realizada uma revisão de bibliografia que visa arregimentar informações e entender mais detalhadamente o assunto para auxiliar na proposição da pesquisa, definição de problemas e objetivos.

Além disso, as pesquisas bibliográficas podem possuir como objetivo verificar o estágio teórico em que um assunto se encontra no momento atual com o propósito de levantar novas abordagens, visões, aplicações e atualizações a partir de material já publicado com o intuito de colocar o pesquisador em contato direto com todo material já escrito sobre o assunto da pesquisa (PRODANOV, 2013).

Desse modo, o primeiro passo desta pesquisa foi a realização de um levantamento acerca da literatura existente referente à desinformação, atores não biológicos e máquinas sociais, inicialmente em uma perspectiva trazida pela Ciência da Informação, entretanto para a definição de algumas das entidades não humanas, foi necessária a busca de literatura de outras áreas. Posteriormente foi pesquisada a literatura que trata sobre a relação entre desinformação e atores virtuais, entretanto, se observou que grande parte deste debate era fomentado pela imprensa e mídias especializadas, ainda sem um olhar mais aprofundado na relação entre ambos.

Não foi realizado nenhum tipo de sistematização nesta busca pela bibliografia utilizada, sendo esta inicialmente concebida na BRAPCI¹⁴, verificando-se, entretanto, que não havia menções sobre questões da desinformação apoiada/realizada por qualquer tipo de entidade não biológica.

Desta forma, a busca passou a ser realizada na base de dados *Google Acadêmico* onde foi encontrado, a princípio, o artigo da Burkhardt (2018) que serviu de ponto de partida para o reconhecimento de outras bibliografias que contribuíram para a pesquisa em maior ou menor grau, surgindo assim o processo conhecido como bola de neve (*Snowballing*) tanto para quem o artigo original referencia, quanto por quem ele é referenciado. A partir dos artigos descobertos também foi realizado o mesmo processo bola de neve até a obtenção de um *corpus* de cerca de 20 artigos.

Para a seção três desta dissertação, que se refere as entidades não humanas, a busca foi realizada também de maneira *ad-hoc*, mas foram encontrados materiais na BRAPCI sobre as nomenclaturas mais comuns (robôs, ciborgues, agentes de *software*) e também foram conduzidas buscas no *Google Acadêmico* para a complementação.

Esta pesquisa bibliográfica resultou nas seções 2 e 3 desta dissertação, desinformação e entidades não humanas respectivamente, e em ambos foi necessário ir além da mera reunião de conceitos e realizar uma classificação e análise para melhor direcionar a pesquisa sobre quais aspectos de desinformação e entidades não humanas seriam relevantes para o estudo.

Quanto ao tema desinformação, se fez necessário observá-la como objeto de estudo e posicioná-la em relação a conceitos correlatos, tais como notícias falsas e pós-verdade. Neste momento foi observada nas formas em que a desinformação pode ser realizada (Subseção 2.2). Ressalta-se que as sete manifestações de desinformação presentes nesta dissertação não representam a totalidade de maneiras que observamos desinformação e algumas das categorias apresentadas ainda carecem de um maior aprofundamento, tais como *Deepfake* e *Data Flood*.

Em seguida se fez necessário observar a natureza da desinformação, uma vez que existe uma tendência de que a desinformação é algo que resulta de um processo enganoso, proposital, com fins prejudiciais e se utilizando de informação falsa. Assim, ao

¹⁴ Esta pesquisa foi realizada entre os meses de março e julho de 2018.

observar no texto de Buckland (1991), também foi percebida a necessidade de apresentar a desinformação sob certas perspectivas, excludentes embora correlatas, que pudessem mostrar que desinformar também pode, em algumas circunstâncias, ser natural, benéfico, altruísta e se utilizando de informação de alta qualidade. E assim como Buckland, neste trabalho, é sugerido que a desinformação pode ser observada como coisa e processo, e um pouco diferente do autor supracitado, neste estudo preferi u-se analisar a informação como propósito, ao invés de conhecimento ou desconhecimento, pois não ficou tão clara esta relação entre desinformação e desconhecimento, uma vez que havia uma riqueza maior de discussão sobre o propósitos e intenções ao desinformar, enquanto que a relação entre desinformação e (des)conhecimento ainda não está (verbo no presente) clara.

Por fim, também era importante apresentar o impacto da desinformação em estruturas sociais e como nós passamos a lidar com a desinformação. E por isso na Subseção 2.3 se observam os movimentos, benéficos e maléficos, que se formaram por conta da desinformação.

Quanto ao estudo das entidades não humanas e máquinas sociais seria inicialmente uma pesquisa apenas sobre máquinas sociais, entretanto, houve dificuldade em posicionar este objeto, em uma visão adequada para a Ciência da Informação e pela maneira como seria preciso apresentar tais máquinas como seres autônomos, com capacidade de interagir com seres humanos e processar informação de modo a influenciar em nossas ações. Desta forma, foi decidido apresentar outros elementos não humanos, onde a maioria é bem conhecida pela CI, e realizando o comparativo com outras sete entidades não humanas (Subseções 3.1 a 3.7), foram posicionadas as máquinas sociais em um contexto mais familiar a CI.

Ainda quanto aos meios, esta pesquisa também se caracteriza como **pesquisa de campo** que, segundo Michel (2009), trata-se da coleta de dados em ambiente natural para que seja possível analisar a vida real, com base em uma teoria, e como esta mesma teoria se encaixa na realidade observada. Segundo a mesma autora, esta pesquisa é apropriada para entender grupos e comunidades, bem como explicar fenômenos, entender realidades e criar significados sociais dos mesmos. Ao investigar a relação da

desinformação e máquinas sociais no processo político Brasileiro trazem à tona este caráter de pesquisa de campo ao trabalho.

Para finalizar a caracterização quantos aos meios, esta pesquisa também pode ser considerada **empírica**. Michel (2009) afirma que a pesquisa é voltada para a experimentação vivenciada e observada dos fenômenos a partir da manipulação de dados, fatos concretos, traduzindo os resultados em dimensões mensuráveis. Ao considerar uma análise de caráter quantitativo para identificar uma relação de “causa-efeito” entre máquinas sociais que se utilizam de desinformação para manipular a opinião das pessoas, tem-se o realce do caráter empírico da pesquisa.

Quanto aos fins, esta pesquisa é considerada uma **pesquisa exploratória**, posto que o estudo realizado discorre sobre uma questão ainda pouco conhecida/explorada no meio acadêmico. Segundo Michel (2009, p. 40) esse tipo de pesquisa “se caracteriza pela busca, recorrendo a documentos, de uma resposta a uma dúvida, uma lacuna do conhecimento”. Portanto, esta observação se propõe a observar/fazer relações e conexões que o ambiente exerce sobre os fenômenos. Neste tipo de pesquisa não há interferência no ambiente e o objetivo desta é explicar os fatos relacionando-os ao ambiente.

Quanto aos métodos de pesquisa utilizados, novamente se ressalta que foram usados diversos métodos para se encontrar evidências que se complementam e trazem um panorama mais completo do papel de tais máquinas sociais na desinformação. O primeiro método a ser elencado para esta pesquisa foi o **estudo de caso**. O estudo de caso se tornou útil neste trabalho pelo fato dele, de acordo com Michael (2009), consistir na investigação de casos isolados ou de pequenos grupos, com propósito básico de entender fatos e fenômenos sociais.

O estudo de caso possui objetivo, metodologia e resultados próprios e se encontra no contexto em que se envolveram usuários virtuais e desinformação nas Eleições Brasileiras de 2018. Os dados colhidos vieram de outros estudos, realizados por outros pesquisadores e não estão padronizados o suficiente para se realizar uma meta-análise.

Para tanto, foram identificados trabalhos originais (fontes primárias) que traziam dados quantitativos sobre a participação de entidades de *software* em processos de desinformação ligados às eleições de 2014 e 2018, bem como o processo do

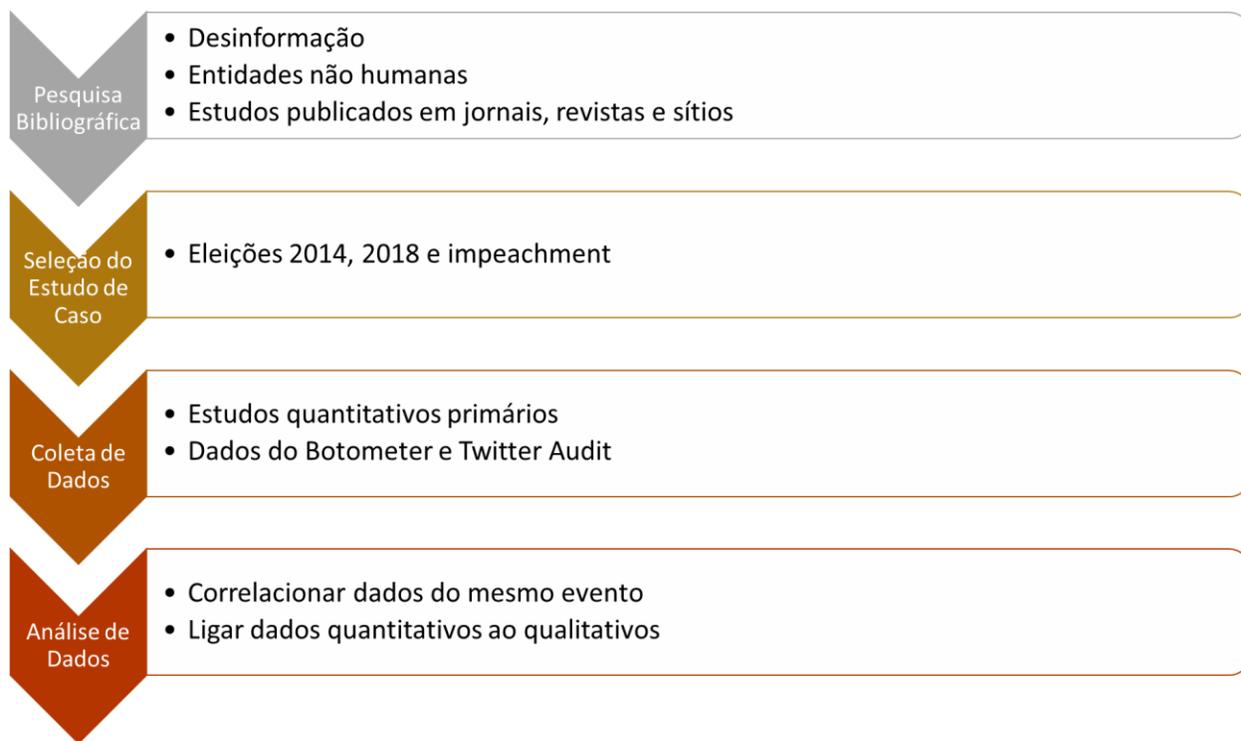
impeachment presidencial em 2016. A intenção inicial era coletar dados apenas quantitativos para a realização de uma meta-análise, mas surpreendentemente havia: (i) um número pequeno de pesquisas quantitativas com o contexto desejado; (ii) os dados publicados não eram padronizados, tornando impossível a aglutinação; e (iii) muitos dados traziam a informação sobre a automação das mensagens ou a informação sobre a desinformação, e estas não vinham correlacionadas, o que exigiu mais análise com o uso de ferramentas próprias para conectar questões específicas.

Por isso, foram também analisados dados qualitativos para preencher as lacunas deixadas pela análise quantitativa agora baseada em dados imperfeitos para a realização de uma meta-análise, mas que, ainda eram bastante ricos para uma análise *ad-hoc*.

4.1 Execução da Pesquisa

Para a realização desta pesquisa foi elaborado um processo apresentado na Figura 4.1 a seguir:

Figura 4 - Etapas da Pesquisa



Fonte: Autoria própria

O resultado da pesquisa bibliográfica resumida para esta pesquisa está presente nas seções 2 a 3 deste documento e já foi comentada na seção anterior, cabendo então ainda explicar como se deu os outros 3 passos restantes.

4.1.1 Seleção do Estudo de Caso

Para se desenhar um panorama mais amplo do papel das máquinas sociais no processo de desinformação foram buscadas situações reais onde tais interagentes foram identificados, em determinados regimes de informação, ricos na propagação ou existência da desinformação em si. A busca pelos “casos de vida real” é mais interessante, uma vez que é possível observar o fenômeno em um estado puro e sem filtro, de seu funcionamento, causas e consequências.

Desta forma, foram buscadas situações que seguiam as seguintes condições:

1. Apelo da imprensa, veículos especializados em tecnologia e/ou acadêmico;
2. Evidências da participação de atores não humanos e que houve algum processo de desinformação, mesmo que estes elementos não estivessem relacionados;
3. Dados quantitativos e qualitativos que possibilitassem criar uma base para se conduzir algum tipo de análise;
4. Mais de uma fonte de informação fornecem estes dados (este ponto é desejável, mas pode não ser estritamente respeitado);
5. Explicação detalhada do processo de desinformação (qualitativo).

Considerando estes cinco critérios, foram vislumbradas três possibilidades de estudos casos:

1. **O processo das eleições Brasileiras de 2014, *impeachment* e eleições Brasileiras de 2018.** Neste estudo de caso seria observado a utilização de máquinas sociais durante o período de 2013 a 2018 e como se deu a evolução do uso, estratégias para atingir os eleitores e opinião

pública, acompanhada de números. O que se observa aqui é uma mudança do uso destes para obter resultados diferentes, ou seja, houve uma evolução não só da tecnologia envolvida, mas do processo de desinformar.

2. **O assassinato da vereadora Marielle Franco.** Neste caso seria observado as postagens relativas ao assassinato da vereadora, dos boatos lançados nas redes sociais e também as estratégias de contra-ataque (vigilantismo) para desmentir os boatos.
3. **Vandalismo na Wikipédia.** Neste caso seria observada a ação de máquinas para reparar a ação de vandalismo nas páginas da Wikipédia. O que cabe observar neste caso é a ação de entidades de *software* modificando conteúdo adicionado por humanos, isso implica em uma capacidade de julgamento e avaliação por parte de tais máquinas.

Ressalta-se que apenas o primeiro foi realizado uma vez que os estudos quantitativos primários sobre o caso da Vereadora Marielle Franco ainda são mais escassos do que os relativos as eleições e este não traria, na avaliação deste trabalho, conclusões muito diferentes das encontradas no estudo de caso do cenário político brasileiro.

Já as questões de vandalismo na Wikipedia possuem uma quantidade considerável de dados quantitativos disponibilizados, entretanto, não há nenhum material qualitativo. Isso exigiria uma análise mais aprofundada e o uso mais rigoroso da meta-análise e por questões de tempo não foram trazidas para esta dissertação.

4.1.2 Coleta de Dados

Os dados coletados são oriundos de duas naturezas: (i) de outras pesquisas publicadas em veículos reconhecidos seja na academia, seja na imprensa e (ii) dados coletados das ferramentas *Botometer* e *Twitter Audit*.

Os dados quantitativos foram coletados a partir de estudos especializados da Fundação Getúlio Vargas (FGV), Laboratório de estudos sobre Imagem e Cibercultura

(Labic) da Universidade Federal do Espírito Santo (UFES), Jornal Estado de São Paulo (Estado), Revista Exame, entre outros. Estes estudos trazem dados quantitativos referentes ao uso de perfis automatizados em postagens e foram coletados entre maio de 2015 e junho de 2019.

Em seguida foi observado que para o evento do *impeachment* os dados qualitativos e quantitativos não eram convergentes, ou seja, ou havia a informação da automação, ou havia a informação sobre a desinformação, mas elas não apareciam juntas. Neste sentido, foi necessário observar nos casos de desinformação se os perfis que enviaram a mensagem eram ou não automatizados (uma das características de entidades não humanas). E por isso foram usadas as ferramentas *Botometer* e *Twitter Audit* para averiguar se os perfis eram ou não utilizados por pessoas reais.

4.1.3 Análise de Dados

Os dados oriundos dos estudos primários foram analisados a partir da relação entre as postagens realizadas por perfis automatizados e a desinformação gerada por estes. Aqui, ressalta-se que os estudos normalmente traziam um panorama geral das postagens, não se restringindo a postagens realizadas por entidades não humanas, o que representa, por vezes, que uma pequena parte do estudo pode ser aproveitada.

Para o caso do *impeachment*, não ficou claro a ligação dos perfis automatizados com a desinformação. Por isso, foi necessário buscar na publicação original os perfis e as mensagens citadas, e foi verificado se os perfis envolvidos eram ou não automatizados com a ajuda das ferramentas *Botometer* e *Twitter Audit*.

O *Botometer* verifica a probabilidade de um perfil ser automatizado. Para isso ele analisa qualquer perfil do *Twitter* observando algumas características das postagens das contas, por exemplo, se a conta posta dois *tweets* no mesmo segundo, ou se a conta posta vários *tweets* em um intervalo pequeno, quantos idiomas o usuário já emitiu postagens, dentre outros comportamentos que indicam alguma conduta “não natural”. Para isso ele atribui uma nota de 0.0 a 5.0 e quanto mais baixa é a nota menor é a chance de o usuário ser um robô. Segundo a própria ferramenta, uma nota entre 4.0 e 5.0 indica que a conta não deve pertencer a um ser humano, enquanto que entre 3.0 e 4.0 significa

que se refere a um ciborgue, uma vez que parte das postagens são automatizadas e outra parte feita por comportamentos mais compatíveis com humanos. Para esta pesquisa, considera-se como perfis automatizados aqueles que têm uma nota 4.5 ou acima.

Já o *Twitter Audit* aponta quantos seguidores do perfil são potencialmente automatizados e quantos se comportam como humanos. Assim ele avalia a proporção entre perfis automatizados e totais e promove uma nota de 0% a 100%, quanto mais alta for a porcentagem mais seguidores “reais” seguem aquele perfil. A lógica do *Twitter Audit* é que estas contas automatizadas costumam seguir e *retweetar* umas às outras, formando uma espécie de *botnet*, quando há uma quantidade alta de perfis automatizados seguindo um outro perfil, há uma chance de o mesmo ser não humano, ou de que o conteúdo seja retweetado por entidades não humanas.

Então, o uso destas ferramentas seria para diagnosticar não só se o perfil que publica a postagem é automatizado, mas se também o mesmo faz parte de alguma *botnet*.

5 DESINFORMAÇÃO NO CENÁRIO POLÍTICO DO BRASIL (2014-2018)

O presente recorte apresenta as evidências da participação de máquinas sociais no processo de desinformação no cenário político-eleitoral brasileiro entre 2014 e 2018. De modo geral é possível verificar o crescimento do uso dessa tecnologia de maneira ainda “não profissional” por ambos os lados que impulsionaram o discurso de ódio ligado ao posicionamento político culminando em uma campanha maciça de desinformação que, a princípio, fomentou-se pelas notícias falsas, posteriormente evoluindo para outras formas de manipulação.

Os acontecimentos aqui abordados compreendem as eleições de 2014, o *impeachment* de 2016 e a eleição de 2018. Em cada um desses eventos foram observados estudos específicos que foram utilizados como fonte de dados para uma análise do contexto e interpretação desses fatos em uma perspectiva cronológica apontando os atores e resultados dessas ações.

Observou-se que o uso de máquinas sociais ocorre desde 2014, porém, somente a partir de 2015, elas se engajam em propagar conteúdo falso. A automação das mensagens foi sofisticada entre 2014 e 2018, na qual as máquinas sociais deixaram de ser atores que replicavam posts indiscriminadamente (*spam*) para se tornarem catalisadores dos usuários com a intenção de criar ou expandir novas bolhas de desinformação.

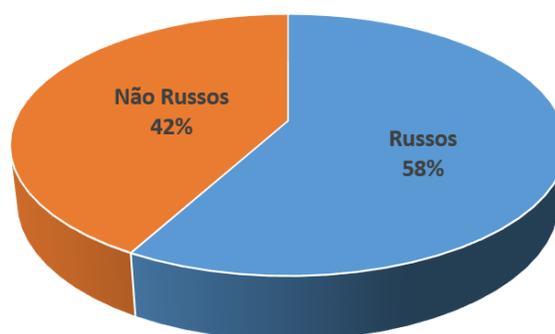
Por fim, pode-se observar que o mesmo “exército de máquinas sociais” ativo no Brasil também esteve envolvido em processos eleitorais em todo o mundo, colocando o Brasil como mais uma peça nessa rede de desinformação, da qual não se sabe quem está envolvido e qual é o seu propósito.

5.1 Eleições 2014

Em 2014, o Brasil entrou no contexto global de atividades de máquinas sociais relacionadas a eleições. Nesse toar, é possível pontuar que em 2018, a empresa *Cambridge Analytica*, que esteve envolvida no escândalo relacionado às eleições americanas e ao Brexit, passou a considerar a possibilidade de abrir um escritório no

Brasil para atuar na eleição de 2018. Porém, cabe ressaltar que a participação de contas automatizadas em eleições brasileiras já havia sido iniciada em 2014, corrobora essa afirmativa o estudo realizado pela Fundação Getúlio Vargas (2018a) que aponta que o candidato de direita, Aécio Neves (PSDB), foi apoiado por 699 perfis, potencialmente não humanos no *Twitter*. Deste total, quatrocentos e quatro (404), cerca de 58%, apresentaram em suas informações pessoais dados muito peculiares da Rússia.

Figura 5 - Origem de contas suspeitas apoiando Aécio Neves



Fonte: FGV (2018a).

5.1.1 Exército de Máquinas Sociais

O mesmo relatório da FGV (2018a) mostrou que os três candidatos com maior intenção de voto usaram contas automatizadas para impulsionar publicações no *Twitter*. Aécio Neves (direita) e Dilma Rousseff (esquerda) usaram, respectivamente, 699 e 509 perfis, administrados por máquinas sociais, no *Twitter*. A candidata Marina Silva (centro-esquerda) não teve o número de perfis automatizados divulgado, mas foi identificado que ela usou a mesma *botnet* que Aécio Neves. Esse conjunto de Máquinas Sociais utilizado por Aécio Neves possuía 508 perfis criados em um intervalo de 2 dias (2 e 3 de agosto de 2013) e publicaram seu último *tweet* entre 5 de outubro e 2 de novembro de 2014 (período eleitoral).

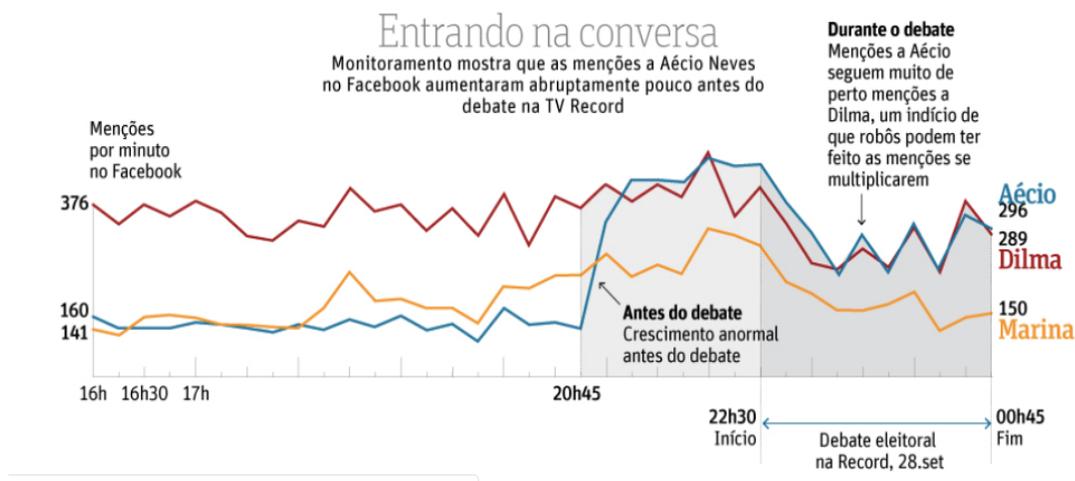
O Jornal Estado de São Paulo (HUPSEL FILHO; GALHARDO, 2015) divulgou uma reportagem apontando que parte dos perfis usados para apoiar o candidato Aécio Neves ainda estariam ativos mesmo após as eleições, algo em torno de 50 perfis (não humanos) estavam propagando material contra a presidente eleita em 2014, Dilma Rousseff.

5.1.2 Participação das Máquinas Sociais

Parece haver um consenso de que a campanha de mídia social de Aécio Neves foi mais eficiente que seus concorrentes no uso de contas automatizadas do *Twitter*. O Jornal Estado de São Paulo (HUPSEL FILHO; GALHARDO, 2015) apresenta um relatório apontando que as mensagens de Dilma Rousseff, no *Twitter*, atingiram 22 milhões de usuários, enquanto a campanha de Aécio Neves alcançou 80 milhões. O estudo da FGV (2018a) mostra que 699 máquinas sociais usadas por Aécio Neves realizaram 773.703 posts no *Twitter* entre os dias 5 de outubro e 2 de novembro de 2014. No mesmo período, os 509 perfis automatizados que apoiaram Dilma Rousseff, postaram apenas 4.695.

Outra pesquisa realizada pelo LABIC (2014) investigou o *Facebook* para mensurar as menções referentes aos três principais candidatos no dia do primeiro debate presidencial, realizado em 28 de setembro de 2014, promovido pela Rede Record de Televisão. A Figura 6 apresenta uma linha do tempo das menções feitas neste dia.

Figura 6 - O Facebook menciona em 28 de setembro de 2014

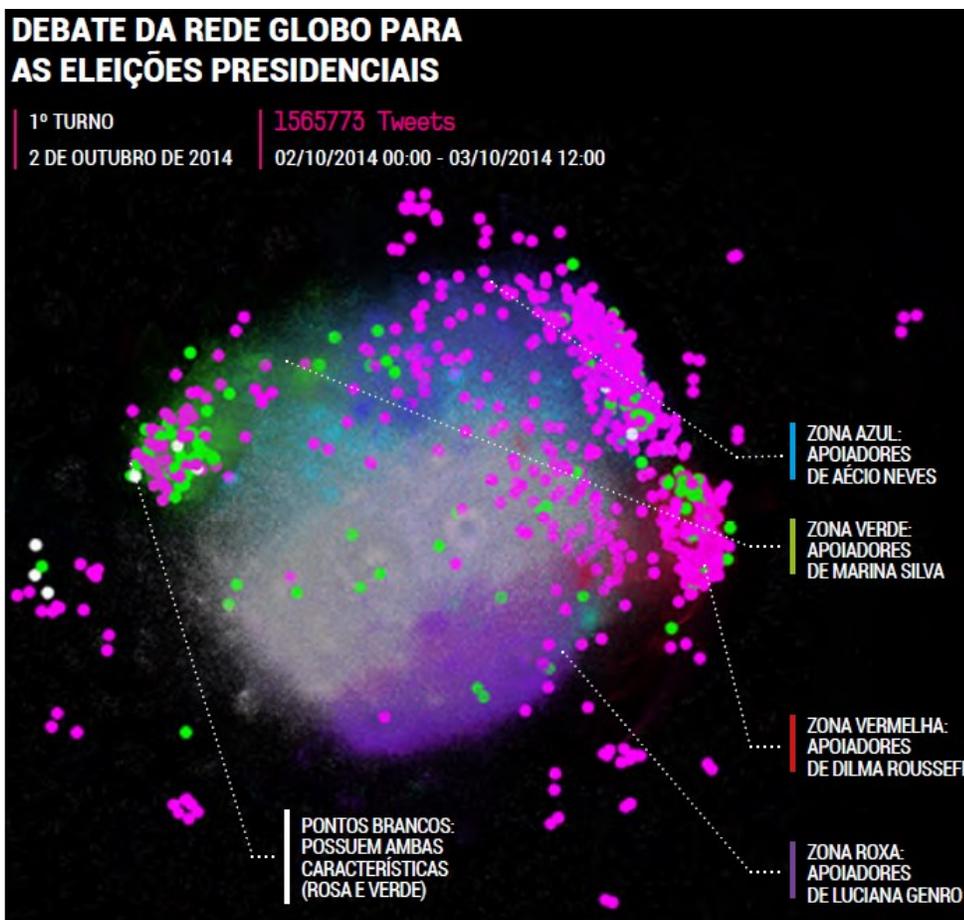


Fonte: Labic (2018).

É possível observar que durante o dia inteiro houve um número de menções regulares a candidata Dilma Rousseff enquanto que para o candidato Aécio Neves as menções quase que triplicaram duas horas antes do início do debate, ultrapassando as menções aos outros candidatos. Durante o debate, as menções entre os candidatos foram em uma quantidade semelhante, o que, segundo o LABIC (2018), indica o uso de contas automatizadas nestes horários.

As menções no *Twitter* referentes a outro debate promovido pela Rede Globo de Televisão, no dia 2 de outubro de 2014, foram monitoradas pela Fundação Getúlio Vargas (2017). Um total de 1.565.773 *twitters* relacionados ao debate foram gerados entre às 00:h00min do dia 02/10/2014 até às 12h00min do dia 03/10/2014 (FGV, 2017). Neste debate estavam presentes quatro candidatos: Aécio Neves, Dilma Rousseff, Marina Silva e Luciana Genro. A Figura 7 resume os resultados do relatório.

Figura 7 - Monitoramento do Twitter durante o Debate da Globo (1º Turno)



Fonte: FGV (2017)

Na Figura 7 pode-se ver um gráfico dividido em 5 clusters. O grupo azul representa usuários que apoiam o candidato Aécio Neves, o grupo vermelho representa os partidários de Dilma Rousseff, o grupo verde representa os apoiadores de Marina Silva, o grupo roxo representa os apoiadores de Luciana Genro e, finalmente, o grupo cinza descreve os usuários neutros.

Pontos rosa representam contas duvidosas que usam algum gerador de conteúdo, os pontos verdes representam contas que postaram pelo menos duas vezes dois *tweets* em menos de um segundo, e pontos brancos representam usuários que possuem as características de pontos rosa e verde simultaneamente. Segundo a FGV (2017), cerca de 98,5 mil mensagens, 6,29% do total, foram geradas por máquinas sociais e a grande maioria delas estava vinculada a algum dos candidatos. Dos *tweets* de apoio ao candidato Aécio Neves, 19,18% foram postados por máquinas sociais. Em favor da candidata Dilma Rousseff, 17,19% das mensagens foram compartilhadas por contas

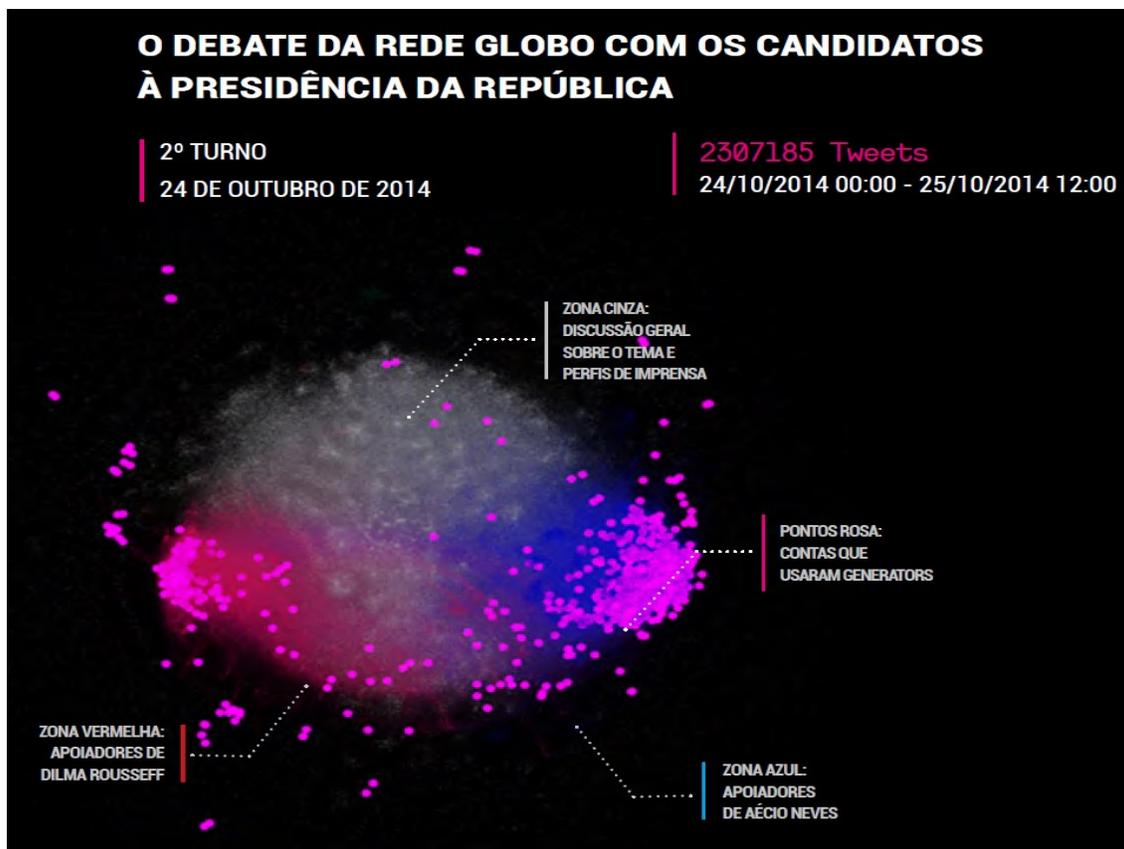
suspeitas. Já Marina Silva e Luciana Genro tiveram 13,52% e 1,64%, respectivamente, de mensagens postadas por estas máquinas.

Em 24 de Outubro de 2014, a mesma emissora promoveu um debate referente ao segundo turno que foi disputado por dois candidatos, Dilma Rousseff e Aécio Neves. Entre as 00:00 de 24 de Outubro de 2014 e 12:00 de 25 de Outubro de 2014 foram coletados 2.307.185 tweets e uma análise semelhante ao debate do primeiro turno foi realizada, como pode ser visto na Figura 5.3.

Nesse debate, as máquinas sociais que apoiaram Aécio Neves foram responsáveis por **19,41% dos posts** enquanto as contas automatizadas fizeram **9,76% das mensagens** em favor da candidata Dilma Rousseff. Assim como na primeira rodada, a maioria das contas duvidosas estava nos polos (azul ou vermelho) indicando o apoio a um dos candidatos. Não foi verificado a presença das máquinas sociais na zona neutra (cinza).

Percebe-se que em três semanas decorridas entre os dois debates, intensificou-se o uso de perfis automatizados, o que sugere a percepção das vantagens de utilizá-los. Se considerar-se a natureza embrionária da adoção desses perfis em disputas eleitorais, é possível notar que a participação do total de cerca de 10% no *Twitter* é significativa e altamente capaz de influenciar os eleitores.

Figura 8 - Monitoramento do Twitter durante o Debate da Globo (2º Turno)



Fonte: FGV (2017).

5.1.3 Desinformação (*hoax*)

As eleições de 2014 não foram marcadas por notícias falsas em si, no entanto, isso não significa que os eventos de desinformação não ocorreram neste período. Alguns rumores (*hoaxes*) foram criados com a intenção de dirimir a imagem dos candidatos. **Os eleitores brasileiros, na época, estavam mais preocupados com a automação da disseminação de mensagens (*spam*) e menos preocupados com sua exposição à desinformação.**

A Revista Exame (2015) apontou os sete principais rumores divulgados relacionados às eleições, dos quais o Tribunal Superior Eleitoral (TSE) e a Polícia Federal investigaram formalmente sobre o candidato Aécio Neves, onde foi publicado que ele poderia encerrar o programa social "Bolsa Família", caso fosse eleito¹⁵. Afora

¹⁵ <http://goo.gl/tCMuo2>

isso, não houve casos noticiados de utilização em massa de eventos de desinformação. Muito menos, foram utilizadas estratégias elaboradas para alcançar eleitores e fidelizá-los com conteúdo de apoio ao candidato de preferência.

5.1.4 Resumo das eleições de 2014

As eleições presidenciais brasileiras de 2014 são marcadas pelo início do uso das mídias sociais em grande escala e a maioria das pessoas foram informadas sobre os candidatos na Internet em vez dos tradicionais programas de TV. Considerando a desinformação, isso ocorreu em uma escala relativamente pequena, e não houve evidência do uso de máquinas sociais na disseminação de fraudes. Os seres humanos foram responsáveis pela criação e disseminação de conteúdo falso.

A mudança do quadro para 2014 é que os políticos passaram a usar as redes sociais como agentes ativos e a militância política emergiu gerando conteúdo de interesse nos diversos veículos da Internet. Considerando os rumores, eles não foram adequadamente investigados pelas autoridades eleitorais e policiais e essa impunidade pode ter encorajado os políticos a considerar o uso em uma escala maior nas eleições de 2018.

5.2 Impeachment 2016

Em 17 de abril de 2016, ocorreu na Câmara dos Deputados a votação do impeachment da presidente Dilma Rousseff. Máquinas sociais foram usadas para fazer propaganda, a favor e contra, a presidente. Observou-se que, entre as eleições de 2014 e o *impeachment* de 2016, parte da mesma comunidade de contas automatizadas que agiam em favor do candidato Aécio Neves estava sendo utilizada para apoiar o *impeachment*. No entanto, a propaganda institucional foi substituída pelo compartilhamento de conteúdo malicioso sobre a presidente eleita.

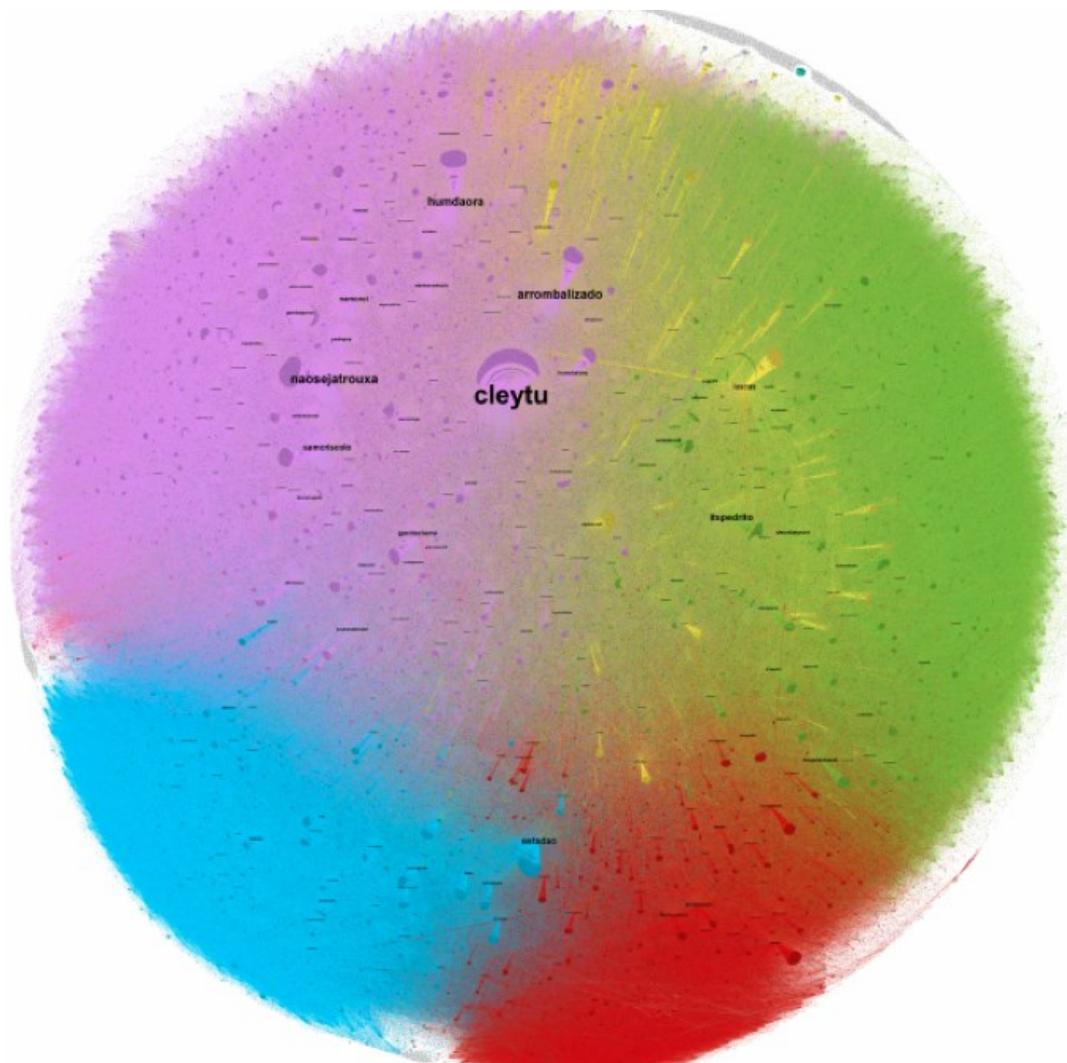
Manifestações em apoio ao *impeachment* aconteceram em março de 2015, apenas três meses após a eleição. A presidente Dilma Rousseff era acusada pelos opositores de cometer "estelionato eleitoral", o que significa que algumas das promessas

de campanha não foram cumpridas. O falso discurso que leva a presidente a uma vitória foi visto por parte dos brasileiros como motivo suficiente para considerar um *impeachment*.

Oliveira *et al.* (2016) apresentam dados das manifestações pós-eleitorais de 2014, em apoio e contra Dilma Rousseff. Os perfis automatizados e usados pelo governo tinham mais alcance e participação no *Twitter* do que aqueles que eram contrários a gestão a época. Malini (2016) apresenta dados do monitoramento do *Twitter* na semana do *impeachment*. Os autores coletaram quase 2,5 milhões de *tweets* contendo as *hashtags* "impeachment" ou "impeachmentday" entre os dias 10 e 17 de abril de 2016. Esse número é muito significativo para as interações no *Twitter* em língua portuguesa, pois, segundo os autores, havia mais *tweets* relacionados ao *impeachment*, em um único dia, do que a copa do mundo de futebol inteira.

Na Figura 9 é apresentado um gráfico resumindo os resultados encontrados por Malini (2016). São 470 mil usuários (nós) conectados por cerca de 2,5 milhões de *retweets* (arestas). Existem cinco áreas principais no gráfico, das quais **roxo, verde e amarelo representam os usuários mais indignados da classe política**. A área **azul representa as pessoas que apoiam o *impeachment***, enquanto a **área vermelha apresenta o apoio ao governo Dilma Rousseff**.

Figura 9 - Atividade no Twitter relacionada ao Impeachment



Fonte: Malini (2016).

O resultado é um pouco surpreendente, porque cerca de 70% dos usuários não estavam focados no debate do *impeachment* em si, mas estavam centrados em críticas ao sistema político brasileiro. O estudo de Malini (2016) não se refere ao desempenho das máquinas sociais nesse período, no entanto, o autor fornece uma lista dos 100 usuários mais influentes no tópico. O autor calculou a influência do usuário usando a fórmula abaixo:

$$\text{Influência} = \frac{\text{Número Total de Menções}}{\text{Número total de tweets}}$$

Com a lista dos usuários mais influentes, decidiu-se verificar se alguns deles são perfis automatizados ou se suas conexões são preenchidas por usuários não humanos. Para tanto foram utilizadas duas ferramentas: *Botometer*¹⁶, que indica a probabilidade de um usuário ser um perfil automatizado (máquina social), e o *Twitteraudit*¹⁷ que avalia se os seguidores do usuário se comportam como perfis automatizados. Essas ferramentas não são totalmente precisas em seus resultados, mas podem ajudar a indicar se as contas são automatizadas. Essa análise foi feita no início de janeiro de 2019 e, para considerar uma conta potencialmente como automatizada, utilizou-se as métricas recomendadas pelo botômetro para um usuário que tenha uma pontuação maior que 4,5 (em uma escala de 0 a 5).

Foi realizada aqui uma pesquisa complementar à pesquisa original de Malini (2016) e fornecer uma dimensão da participação das máquinas sociais neste processo. Dos 100 usuários mais influentes apresentados no estudo original, 28 tiveram pontuação maior que 4,5, o que indica uma alta probabilidade destes perfis não pertencerem a usuários humanos. A rede de seguidores desses usuários também foi avaliada, e 18 deles tinham mais de 15% de seguidores também indicados como contas automatizadas. Estes números indicam, embora não se possa medir o quanto, que máquinas sociais certamente participaram do debate sobre o *impeachment*.

Em outro estudo, a FGV (2017) analisou 826.493 *tweets* coletados em 13 de março de 2016, um mês antes da votação do *impeachment*. O gráfico resultante desta pesquisa pode ser visto na Figura 10. Este estudo corrobora o de Malini (2016). Aqui é possível observar uma grande área cinza, indicando opiniões neutras sobre o *impeachment*. A área azul representa a parte que apoia o *impeachment* e os manifestantes contrários são representados na área em vermelho. Ressalta-se que o levantamento da FGV (2017) usou o *impeachment* como ponto de partida para discutir a política brasileira, sendo esses dados obtidos antes da semana do *impeachment*.

Segundo FGV (2017), pelo menos 10% das interações ocorridas neste dia (13/03/2016) foram feitas por usuários não humanos. Do lado da presidente Dilma

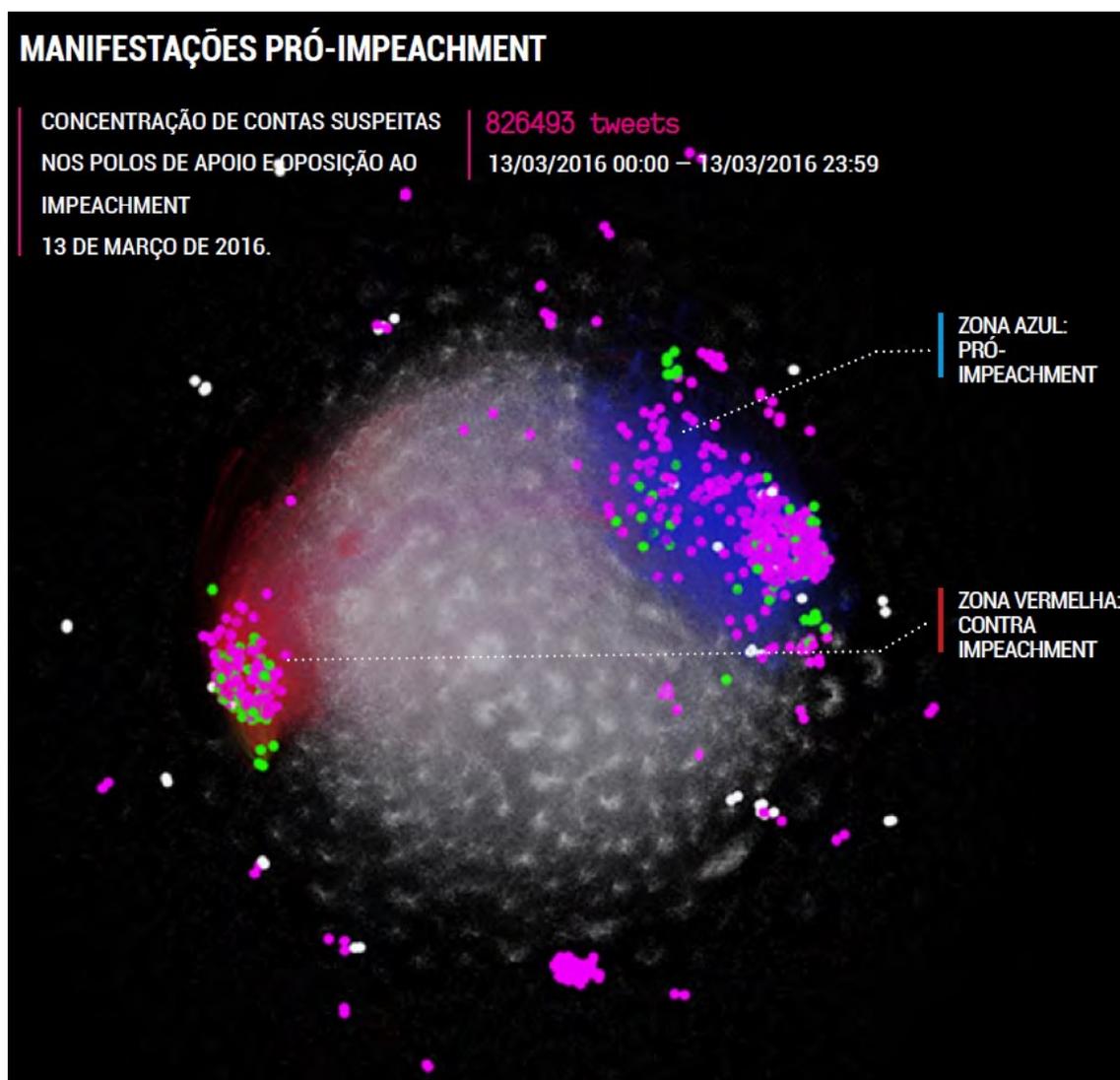
¹⁶ <https://botometer.iuni.iu.edu/#/>

¹⁷ <https://www.twitteraudit.com/>

Rousseff, 21,43% das mensagens de apoio vieram de contas automatizadas, enquanto do lado pró-impeachment, 16,61% das interações foram feitas por perfis falsos.

Na zona cinzenta, poucas máquinas sociais foram detectadas. Arnaudo (2017) afirma que o uso destas entidades artificiais, durante os protestos relacionados ao *Impeachment* da presidente Dilma Rousseff, corroborou para a divulgação de notícias falsas. A de maior destaque é aquela que afirmava que o Partido dos Trabalhadores (PT) implantaria um regime comunista no Brasil, informação que foi credenciada por 64% dos brasileiros, na época.

Figura 10 - Concentração de Contas Suspeitas (Impeachment)



Fonte: FGV (2017).

Em reportagem apresentada pela BBC (2016) relata-se que na semana de *impeachment*, três das cinco notícias mais compartilhadas no *Facebook* eram falsas. Segundo o artigo, cerca de 200.000 pessoas receberam esse conteúdo, eles avaliaram 8.290 notícias de 117 jornais, revistas e blogs e resultaram em 6,1 milhões de aquisições. Essas três notícias, sozinhas, foram compartilhadas por 115 mil vezes.

5.2.1 Resumo do Impeachment

Em 2016, a impressão de que as eleições de 2014 ainda não havia terminado permanecia. O candidato derrotado manteve parte de sua comunidade de perfis automatizados ativa e mudou sua estratégia de promover ideias próprias para deteriorar a imagem de oponentes ao postar notícias sobre escândalos, algumas das quais falsas. Diferentemente das eleições de 2014, onde houve alguns boatos sobre os candidatos, foi possível perceber uma estrutura organizada para a proliferação da desinformação, onde grupos independentes geravam conteúdo falso, que era replicado por máquinas sociais. Por fim, é importante notar que persiste a sensação de que grande parte da população estava mais envolvida em demonstrar sua insatisfação com a classe política do que em apoiar um grupo político específico na questão do *impeachment*.

5.3 Eleições 2018

As eleições presidenciais brasileiras de outubro de 2018 também foram marcadas por notícias falsas e o uso de máquinas sociais, e a estratégia de usar grupos independentes para gerar conteúdo malicioso disseminado por perfis automatizados foi reforçada.

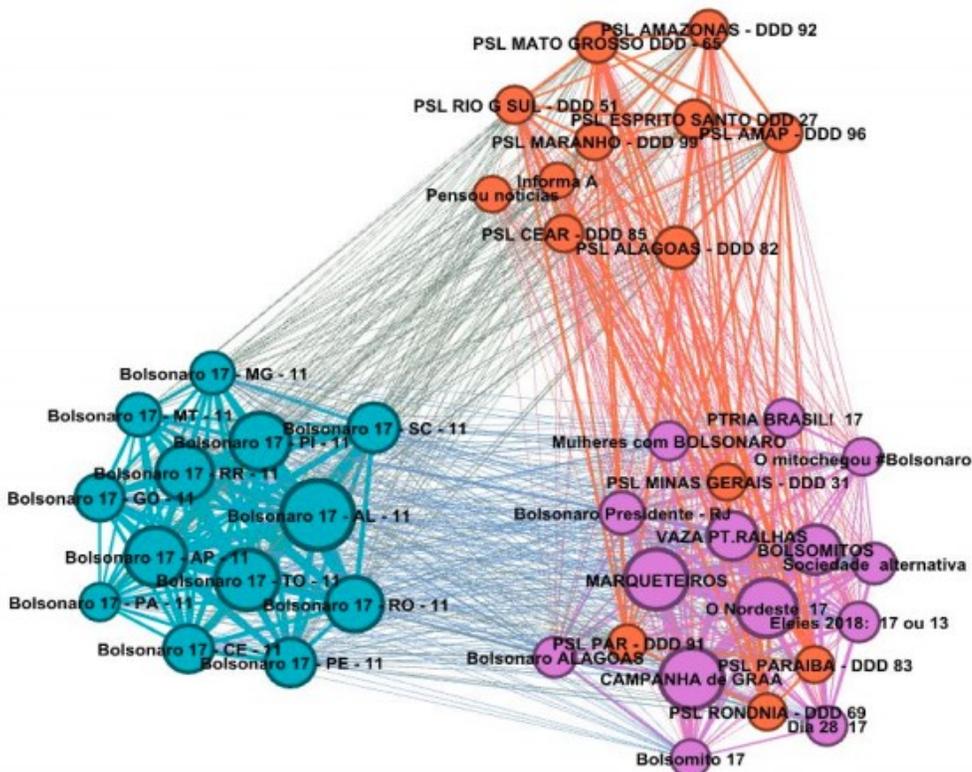
A Agência Lupa (2018) avaliou 846.905 mensagens do *Whatsapp* de 347 grupos relacionados a políticas públicas. Destas, 107.256 eram imagens, e deste universo, as 50 imagens mais compartilhadas foram analisadas, e apenas quatro eram totalmente verdadeiras. Das 46 restantes, 9 foram classificadas pela agência como teorias conspiratórias, 8 foram consideradas falsas, 7 apresentavam conteúdo fora do contexto, em 6 imagens eram expostas opiniões, 6 apresentavam conteúdo exagerado, 4 se

tratavam de rumores, 3 traziam sátiras e, finalmente, 3 traziam fatos que não puderam ser verificados. Essas imagens foram reproduzidas cerca de 3000 vezes e a Agência Lupa não divulgou o número total de usuários que receberam essas mensagens.

Em outra pesquisa promovida pelo Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS-RJ), foram observados 110 grupos relacionados a política apoiando o candidato de Jair Bolsonaro no *Whatsapp*, no período de 17 de outubro a 23 de outubro, entre o primeiro e o segundo turnos das eleições de 2018. O ITS (2019) observou 26.487 mensagens, das quais 2.242, próximo a 10%, foram enviadas por apenas 10 usuários e, devido ao comportamento desses usuários, há uma considerável probabilidade de serem máquinas sociais.

O ITS (2019) também sugere que foram formadas bolhas de interconexão de usuário em vários grupos, criando uma sobreposição de influenciadores em cada um deles. As máquinas sociais não ficam isoladas em um único grupo, e muito menos, cada grupo tem um único perfil não humano. A figura 5.7 mostra a interconexão nesses grupos.

Figura 11 - Interconexões de usuários no grupo de eleições do WhatsApp



Fonte: ITS (2018)

Cada nó representa um grupo e seu tamanho varia de acordo com o número de membros compartilhados com outros grupos específicos. A borda mais espessa significa que mais usuários são compartilhados entre os grupos conectados. A teoria das bolhas é reforçada quando se observa que muitos grupos possuem o mesmo administrador e os usuários mais participativos são, provavelmente, as máquinas sociais.

Em outro estudo, Machado *et al.* (2018) analisaram 1.432.000 de *tweets* oriundos de 204.097 usuários do *Twitter* entre 19 de agosto e 28 de agosto de 2018. Os autores afirmam que apenas 2% das ações do *Twitter* se referem a "*junk news*" e são feitas por uma pequena parcela de usuários. Cerca de 80% do conteúdo falso favorável ao candidato Jair Bolsonaro (direita) veio de um pequeno grupo de militantes, enquanto que os partidários mais atuantes de Lula e Haddad (ala esquerda) foram responsáveis por 65% da disseminação de conteúdo falso. Os autores concluem que a rede de notícias indesejadas de Lula e Fernando Haddad publica com mais frequência, mas alcança menos pessoas do que o grupo que apoia Jair Bolsonaro. Repetindo o que se viu em 2014.

Por fim, em um outro estudo, a FGV (2018b) observou 232 contas no *Twitter* que publicaram 8.185 posts sobre a política brasileira em língua portuguesa entre 1 de agosto e 26 de setembro de 2018. O que chama a atenção neste estudo não estão nas mensagens publicadas (*tweets*) em si, mas nos perfis que postam estas mensagens. Suspeita-se que estes tenham sido parte de algumas comunidades de perfis automatizados que está em funcionamento desde 2011 e que também tenham atuado em outras eleições ao redor do mundo. Uma das contas apresentava postagens em mais de 40 idiomas, com conteúdo muito semelhantes entre si.

5.3.1 Resumo das Eleições 2018

Notícias falsas e máquinas sociais continuaram a serem atores relevantes nas eleições de 2018, no entanto, a estratégia de usá-las foi diferenciada em relação as eleições de 2014. Agora, políticos e partidos faziam campanha pela propaganda oficial, enquanto grupos militantes eram responsáveis por criar e disseminar conteúdo falso. O *Twitter* e o *Facebook*, que são mais fáceis de monitorar, trouxeram uma quantidade menor de material falso, enquanto o *Whatsapp* se tornou a mídia preferida para conduzir a campanha de desinformação. A estratégia, iniciada no período de *impeachment*, baseada em grupos militantes que geram conteúdo malicioso disseminado por máquinas sociais foi mantida, mas aparentemente essas ações foram migradas para o *Whatsapp* onde há maior privacidade no compartilhamento de dados, e a relação entre as pessoas tende a estar mais próxima.

6 CONCLUSÃO

Uma vez concluída a pesquisa, se faz necessário destacar as principais contribuições e avanços trazidos à tona e, neste sentido, esta conclusão será dividida em cinco sessões que tratam da descobertas, discussão dos resultados, ameaças a validade, impactos e trabalhos futuros.

6.1 Descobertas

Após analisar o estudo de caso, pode-se elencar que a relação entre máquinas sociais e a desinformação no contexto político-eleitoral brasileiro nos anos de 2014 a 2018 trazem as seguintes características:

Conteúdo falso não é maioria: Sem dúvida, o conteúdo falso tende a ser mais difundido, e as máquinas sociais são atores relevantes nesse processo. Mas vale destacar que, apesar da imensa quantidade de conteúdo malicioso disseminado nas redes, há um volume ainda maior de propaganda oficial e verdadeira propagada pelos próprios candidatos e que são replicadas também por perfis automatizados. Como a grande maioria do eleitorado age com menos paixão, eles tendem a compartilhar notícias divulgadas pela mídia oficial dos candidatos ou veiculados na imprensa, o que minimiza o risco de compartilhar a desinformação.

Existe o trabalho da militância não oficial em utilizar máquinas sociais para difundir conteúdo falso, entretanto, isto ocorre em um volume significativamente menor quando comparado ao conteúdo verídico. Isso significa que, quando se trata do uso de máquinas sociais para campanhas políticas-eleitorais, atualmente, esta se falando da massificação da disseminação de informação ou da criação do burburinho (*buzz*) inicial que ocorre em mídias sociais para viralizar um determinado conteúdo. Assim não se pode associar, ainda, as máquinas sociais puramente à desinformação.

Isso significa que em determinados momentos, não são os perfis automatizados que mantém o ritmo de disseminação de conteúdo inverídico cabendo esta movimentação aos usuários humanos resultando em um ambiente heterogêneo na

disseminação da desinformação. E também significa que perfis automatizados também replicam um volume considerável de propaganda oficial criada por usuários humanos.

Diferentes mídias sociais, diferentes abordagens: O *Facebook* e *Twitter* estão sendo utilizados para realizar a propaganda oficial. Afinal, são nessas plataformas onde os usuários podem realizar algum tipo de curadoria para reduzir a exibição de notícias falsas e bloquear perfis indesejados. Isso significa que estes meios, embora sejam mais utilizados, permitem ao usuário comum algum tipo de controle para filtrar que tipo de mensagens e perfis eles desejam receber.

Nesses meios, apenas canais mais polarizados, criam e compartilham notícias falsas, entretanto elas são passíveis a: (i) um controle maior dos próprios usuários que podem denunciar o perfil tornando inutilizável; (ii) também podem ser monitorados mais facilmente por agentes públicos de regulação, tais como polícia, tribunais eleitorais e ministério público; e (iii) é mais fácil vincular estes perfis e outros perfis oficiais, bem como encontrar com maior facilidade o responsável pela conta, o que facilita em investigações contra futuros crimes e ilegalidades. Neste sentido, existe um certo controle sobre conteúdo nestas redes sociais inibindo o comportamento de quem costuma desinformar.

Já no *Whatsapp*, existe a aproximação entre os usuários, bem como pessoas que não pertencem ao grupo tem uma dificuldade maior de acessar a mensagem original. A familiaridade presente nesta rede propicia o *word of mouth* (boca a boca), que é uma estratégia eficiente no *marketing* onde um amigo recomenda alguém ou algum conteúdo ao usuário, tornando-o mais atraente e de maior credibilidade. Esta espécie de propaganda é eficaz em campanhas eleitorais já que um amigo ou parente recomenda um candidato ou posição política.

O anonimato proporcionado pelo *Whatsapp* a pessoas externas ao grupo dificulta o controle das mensagens por agentes reguladores. Pois, estes agentes precisam participar dos grupos e conhecer o autor das mensagens para que uma investigação possa ser conduzida. Existe, também o fato do *Whatsapp* estar vinculado a número telefônico e não a uma pessoa, possibilitando que usuários de telefones pré-pagos sejam completamente invisíveis, o que incentiva o uso não muito ortodoxo deste canal.

Também é mais fácil criar bolhas em grupos de *Whatsapp* onde apenas conteúdo de apoio a um determinado candidato seja veiculado. A facilidade existente no recurso de “encaminhar” possibilita que usuários sejam bombardeados de mensagens que influenciem os eleitores. Neste sentido, é possível que usuários tenham acesso a grupos que máquinas sociais estejam sendo utilizadas para apoiar um determinado candidato, contribuindo com conteúdo falso.

Máquinas Sociais nunca descançam e as eleições nunca terminam: No final das eleições as atividades relacionadas às máquinas sociais não foram encerradas, isso foi percebido nas eleições de 2014 e também foi verificado nas eleições de 2018¹⁸. Isso significa que tais perfis não só trabalham 24 horas por dia e 7 dias por semana, mas também fora de períodos eleitorais.

O que se observa desde o fim das eleições de 2014 é que um conjunto de perfis automatizados continua trabalhando para disseminar conteúdo, verídico e inverídico, na intenção de diminuir a credibilidade e status da posição política oposta. O processo de mobilização do *impeachment* de 2016 começa poucos dias depois do término da eleição de 2014, ao ponto que, em 2015, já se existam manifestações populares intensas se manifestando pela possibilidade de *impeachment* (favorável ou desfavorável).

Assim se observa que o uso de máquinas sociais para disseminar conteúdo inverídico é apenas uma das táticas usadas pelos políticos para marcar sua posição, mas esta ação participa de uma estratégia maior, e mais constante, de desinformar como processo. Isso implica que criar uma imagem forte do candidato apoiado através de ações afirmativas que tem apoio da população, embora não se trate de notícias falsas, criam uma situação de *data flood* e *Internet bias* favoráveis ao candidato apoiado.

Então, não se deve reduzir a questão da desinformação proporcionada por contas automatizadas apenas no âmbito das notícias falsas e pós-verdades, mas na construção de uma imagem, falsa ou enviesada, dos atores políticos. E isso ocorre no dia a dia e não nas campanhas eleitorais.

Máquinas Sociais fazem parte das bolhas de notícias falsas: Máquinas Sociais são mais eficazes quando participam dos mesmos grupos e quando interagem umas com as outras. Assim, elas não se tornam o único usuário “notável” a compartilhar um

¹⁸ <https://bit.ly/2RhPWmE>

determinado tipo de conteúdo e que não é recebido por ninguém. Quando se vê a interação entre membros de um grupo sobre determinado conteúdo, não apenas se é levado a prestar mais atenção, como também se passa a sentir mais confortável a participar do debate, principalmente, se o usuário percebe que a sua posição será melhor recebida pela maioria dos outros membros.

Quando as comunidades de perfis automatizados agem em conjunto nos mesmos grupos, passam a sensação de que muitos usuários compartilham as mesmas ideias e defendem os mesmos pontos de vista, o que intensifica esta visão e cria uma “falsa” multidão que é capaz de influenciar mais pessoas. Assim, dificilmente, uma máquina social irá estar isolada em um meio de humanos, o mais provável, é que exista uma *botnet* coordenada para tratar da desinformação.

Alcance sobre o volume: Nos eventos das eleições de 2014 e do *impeachment* de 2016 tanto o monitoramento das mídias sociais quanto a preocupação das pessoas foram mais voltadas ao volume de mensagens enviadas. Havia o medo de que intermináveis fontes de *spam* viessem a atormentar os usuários e se creditava às máquinas sociais essa automação. E de fato, nestes dois eventos, os boatos e notícias falsas foram propagados em mídias sociais de massa (*Facebook* e *Twitter*), mas já houve uma resposta da sociedade em filtrar estes conteúdos e punir os responsáveis pela veiculação.

Já em 2018, houve uma maior atenção ao alcance desta comunicação. Atingir mais usuários significa mais chances de encontrar pessoas dispostas a compartilhar conteúdo, criando um efeito de bola de neve. Então não foi necessário criar um volume imenso de conteúdo desinformativo, mas sim elaborar uma estratégia sobre que conteúdo verídico e inverídico poderia trafegar nas redes, e como cada um atingiria os usuários alvo.

Assim, mais do que criar notícias falsas e disseminar na rede, se percebeu, por parte do candidato vencedor, uma estratégia de criar bolhas de usuários e a partir deles compartilhar conteúdo, verídico e inverídico, em estratégias bem definidas de como cada notícia irá atingir que segmento do eleitorado. Então menos notícias atingindo um grupo maior de pessoas se mostrou, mais eficiente, do que uma quantidade grande de postagens disseminadas de forma, aparentemente, viral.

Conteúdo sobre frequência: Em 2014, as ações das máquinas sociais na rede foram testadas em momentos cruciais do processo eleitoral, como dias de debates ou dias próximos às eleições. O que implica que em determinados eventos, foram espalhadas propaganda eleitoral sobre os candidatos. Isso mostrou a clara estratégia comum em eleições de “distribuir santinhos” em comícios, debates e em colégios eleitorais, sendo que no meio virtual. Assim, material de campanha era distribuído de maneira mais frequente em debates ocorridos em meios de comunicação.

Em 2018, foi mais visível a noção de que o conteúdo era compartilhado a qualquer momento; mas o “quê” publicar se tornou mais relevante do que “quando”. Notícias falsas tiveram maior circulação em determinados eventos como quando eram veiculadas notícias ligadas a investigação sobre um determinado candidato ou seus partidários. Este “*timing*” de postar era mais eficiente do que as postagens padrões que ocorriam depois de debates ou aparições públicas.

Assim, colocar a propaganda em um contexto, e conseqüentemente a desinformação neste mesmo contexto, se tornou mais efetivo do que a propaganda convencional, que continuava a existir nos canais oficiais. Sempre que havia algum escândalo ou elogio a um candidato, algum tipo de postagem era veiculada para maximizar o resultado desejado, e muitas dessas envolviam a desinformação.

O Brasil e o mundo pequeno: A comunidade de perfis automatizados que atuou nas eleições brasileiras também esteve presente em outros países, sugerindo que o Brasil faz parte na estratégia mundial de desinformação com o auxílio das Máquinas Sociais. Existe algum tipo de estratégia global para manter/levar alguns grupos políticos ao poder, e o Brasil faz parte deste fenômeno.

Isso coloca o Brasil como um ator globalmente relevante para estes grupos políticos e também se pode observar quem faz parte de cada grupo e como estes agem globalmente.

Máquinas sociais e o efeito bola de neve: Quando se trata de desinformação, as máquinas sociais foram mais utilizadas para iniciar a movimentação das notícias falsas deixando para os humanos fomentar os debates sobre tais notícias, bem como fazer um movimento maior da disseminação deste conteúdo inverídico. Ou seja, as contas automatizadas, disseminam a desinformação para um grupo de pessoas como o

início de uma bola de neve, que vai sendo inflamada por usuários humanos, fazendo com que aquela desinformação seja mantida em evidência por mais tempo e exibida para um grupo maior de pessoas.

Neste sentido, a intenção das máquinas sociais não é monopolizar a disseminação e debate das notícias falsas, mas sim iniciar o movimento, deixando as pessoas analisarem, debaterem e disseminarem (num segundo nível) a desinformação. Assim, é possível perceber que estratégias ligadas a efeito manada e criação de bolhas são as estratégias que mais foram utilizadas em 2018.

Máquinas sociais se tornam mais sofisticadas: A sofisticação observada em 2018 indica que as máquinas sociais não são mais usadas como ferramenta de *spam*, disseminando um grande volume de notícias falsas aleatoriamente, mas sim como um instrumento para encontrar outros usuários que se juntarão à rede criando bolhas. Como o boca-a-boca é mais eficiente do que a propaganda em massa, o comportamento da máquina social se torna mais inteligente, e o uso de notícias falsas é realizado seletivamente nos momentos mais apropriados.

6.2 Discussão dos Resultados

A partir dos resultados obtidos através das análises realizada pelo estudo de caso trazido no presente trabalho, é possível discorrer sobre os pontos identificados na sessão anterior.

Primeiramente insta salientar que as descobertas apresentadas como substrato da arguição efetivada não representam um rol taxativo das características evidenciadas nas ações de desinformação envolvendo máquinas, pelo contrário, o que foi constatado relaciona-se ao estudo de caso em evidência, mas não exclui ou limita o aparecimento de outros atributos em face das peculiaridades que cada caso de desinformação carrega consigo. Desta forma, serão tecidas pontuações relacionadas exclusivamente as 11 (onze) características que foram identificadas no decorrer da pesquisa.

Observando o primeiro aspecto descoberto, **o conteúdo falso não é maioria**, nota-se que apesar de, aparentemente, o maior volume de informações visualizadas em ambientes, cuja há ocorrência de desinformação é verificada seja falso, esse fato não

passa de uma impressão ocasionada pelo poder que as máquinas têm de replicada de maneira agressiva conteúdo em rede.

Apesar disso, se por um lado o volume de informações verídicas é maior do que as informações inverídicas, por outro lado pode-se observar que o impacto da desinformação é maior quando exponencializado por perfis automatizados, em razão da velocidade e capacidade que esses apresentam.

Entretanto, deve-se afastar da ideia de que a utilização dessas ferramentas ocorre apenas para a disseminação de conteúdo com relevância para o ato de desinformar, pois como já pontuado é verificado o emprego dessa tecnologia para disseminação de conteúdo verídico, como por exemplo a propaganda eleitoral.

Contudo deve ser observado que a ausência de medidas que barrem as investidas de desinformação no âmbito das redes sociais, pode encorajar ainda mais o uso dessa técnica, não impedido que em determinado momento o conteúdo falso torne-se maior parcela daquele disseminado por máquinas, ocasionando desdobramentos mais sérios do que os já conhecidos.

É constatado também que **as estratégias desses atores mudam conforme a rede de atuação**. A migração de plataforma para realização de investidas que desencadeassem processos de desinformação é fato que se constata pela simples observância dos períodos estudados na pesquisa.

Nesse contexto, deve ser considerado o deslocamento habitual de usuário de uma plataforma para a outra, o que decorre do surgimento de novas ferramentas em detrimento do desuso das existentes. Contudo, não se pode deixar de lado o fato de que as redes sociais que se expandiram após 2014, como por exemplo o *Whatsapp*, proporcionam maior liberdade e 'anonimato' para quem visa proceder com ações de desinformação, favorecendo também a expansão desse tipo de medida.

Destaca-se que no caso do *Facebook* e *Twitter* é possível identificar que as investidas dos perfis que demandam conteúdo inverídico possuem características marcadamente polarizadas e atividades automatizadas que podem ser mais facilmente identificadas pelo usuário. Isso porque são em geral contas que apenas replicam conteúdo e baixo engajamento quando da produção de conteúdo próprio ou outras ações.

Ou seja, essas duas plataformas permitem um maior controle por parte do usuário, não só do aspecto de quem produz, mas também do que é produzido, realizando assim o que se chama de curadoria do material veiculado nas redes sociais.

Diferente ocorre no caso do Whatsapp, onde verifica-se uma mudança brusca de estratégia no *modus operandi* dessas máquinas e de quem produz conteúdo, principalmente pelo formato como se posiciona a referida plataforma. Depreende-se que, por se tratar de uma rede cuja proximidade e semelhança entre os seus membros seja maior, a facilidade na disseminação de conteúdo é propiciada pelo fator confiança que integra aquelas relações.

Os seja, o rigor das notícias fabricadas precisa observar menos a técnica e mais o apelo imagético e de textos curtos com fácil compreensão para sensibilizar o compartilhamento entre usuários humanos, criando assim as bolhas informacionais e a enxurrada de conteúdo que a princípio é iniciada por máquinas, mas repercutida em grande parte por usuários humanos.

É com base nessa realidade que as empresas responsáveis por essas plataformas iniciam investidas, ainda muito tímidas se comparados aos danos já causados, para conter esse tipo de ação. Por outro lado, se mostram resistentes em abrir mão dos seus dados para apreciação do Estado que busca compreender esses mecanismos e criar formas de prevenir/reprimir o uso dessas ferramentas, principalmente no âmbito da política.

O poder de **adaptação das máquinas conforme o contexto de atuação** se revela como fator que propicia vantagem econômica para quem dispõe desse tipo de tecnologia. Desta forma, ao constatar que elas permanecem ativas mesmo após o termo final dos eventos que estiveram atuantes.

Ou seja, há o investimento único para adquirir um *botnet* que seguirá em plena atividade passando por pequenos ajustes para evitar desvio ou a geração de informações descontextualizadas. Desta forma, além de desinformar, essas máquinas aumentam a visibilidade das informações de quem as detém, visto que a todo instante estará em rede compartilhando o conteúdo para qual foi programada.

Aqui não se está falando necessariamente de conteúdo falso, mas da capacidade de volume de informações que esse tipo de tecnologia poderá disseminar com o intuito

de induzir ou mergulhar o usuário da rede em filtros bolhas, dificultando inclusive a busca por conteúdos distintos em razão da cortina de fumaça que proporciona.

As **máquinas sociais atuam em conjunto** em espaço compartilhado com usuários humanos. Com isso, é possível notar que o engajamento entre máquinas é intencional, apresentando uma semelhança nas ideias apresentadas, trazendo para os participantes da rede a sensação de que muitos compartilham do mesmo pensamento. Isso favorece o engajamento de tantos quantos estiverem em rede, afastando a ideia de que o comportamento daqueles perfis é artificialmente produzido.

O engajamento no compartilhamento de conteúdo também é propiciado por esse aspecto de colaboração entre máquinas, que por si só já replicam informações em quantidades consideráveis, mas a partir do momento que outros perfis (automatizados) também o fazem criam a falsa sensação de credibilidade daquilo que está sendo disseminado, assim encorajando outros usuários a também fazê-lo.

O **alcance e sensibilização** dos usuários (humanos) para que esses interagissem com o conteúdo compartilhado se apresenta como uma das facetas que representa a evolução no uso dessas máquinas, isso porque, se antes era comum apelar apenas para o volume de conteúdo na rede com o intuito de gerar manifestações como a enxurrada de dados, filtros bolhas ou o viés da Internet, hoje esse tipo de ação já não basta, ao passo que é mais facilmente identificada pelo cidadão.

Desta forma, buscou-se investir em estratégias que influenciasses os usuários a também participar da rede de compartilhamento, gerando engajamento de perfis reais nessas ações, criando o efeito bola de neve a partir do 'recrutamento' de pessoas dispostas a disseminar aquele conteúdo.

A questão nesse ponto torna-se mais complexa e artilosa, pois o que a princípio se iniciou de maneira eventual (e talvez despreziosa), tornou-se um mecanismo de manipulação especializada cujo alcance ainda se desconhece.

O **aperfeiçoamento do conteúdo produzido** destacou-se como estratégia utilizada nas últimas eleições vindo a substituir a tática de apelar para distribuição de informações em momentos oportunos, conforme verifica-se em cenários políticos anteriores. Nesse toar, nota-se que a disseminação de informações verídicas ou não se dava com o intuito de promover ou denegrir aos atores envolvidos no contexto. Contudo,

o teor do que era veiculado atendia a uma demanda específica, conforme a situação que naquele momento estava em evidência.

Assim, destacou-se ações que iam além da desinformação, devendo ser considerada manifestações como a Internet *bies*, *data flood*, *social spam* como métodos para tumultuar, distorcer ou abafar as informações veiculadas nas mídias oficiais. Ou seja, verifica-se que não facilitar a busca de informações na rede por parte do usuário (comum) foi também um instrumento para prosperar a campanha de desinformação que se instaurou no decorrer da disputa política.

As máquinas empregadas com o objetivo de desinformar compõem uma **botnet de ordem mundial** que migram por países a medida que é instaurado um contexto eleitoral, elas podem até permanecer ativas passado esse momento, conforme já foi visto, contudo é possível notar também que máquinas atuantes nas eleições estadunidenses também foram atuantes nas eleições brasileiras, indicando que o Brasil é apenas mais um ator que integra esse esquema de escala global.

A sofisticação dessas ferramentas deve ser analisada como fator preponderante para o sucesso e eficácia dessas campanhas que prometem expandir-se de maneira cada vez mais agressiva se medidas de prevenção e contenção não surgirem para embate dessa problemática. Carece de ressaltar que as estratégias que, antes eram primitivas pelo fato das máquinas focarem apenas na massificação de informação, hoje assumem contornos refinados pois consiste em apenas iniciar o debate de maneira favorável a quem as contrata, deixando para os usuários humanos o poder-dever de disseminar aquilo que foi posto.

6.3 Ameaças a Validade

A primeira ameaça a validade identificada é relacionada a decisão de se **realizar um estudo secundário**. A partir do momento em que são utilizados, apenas, os dados analisados e interpretados por outrem, se torna muito restrito o que pode ser feito com aqueles dados, assim, não há como garantir que haja a coerência entre os dados originais e os dados compilados. Portanto, cabe explicitar o reconhecimento de que os

possíveis vieses e falhas existentes nos estudos primários são, de certa forma, carregados para este trabalho.

Para minimizar os problemas foram tomadas algumas medidas. Inicialmente, vale considerar que tudo o que foi utilizado foi visto e revisto para evitar que se criasse também vieses e erros oriundos da interpretação dos dados. Todos os gráficos e textos foram conferidos e as informações foram cruzadas e checadas; quando se tratava de dados brutos, também foi realizada pela autora o processamento/cálculo sugerido na fonte original como forma de validar os dados lá apresentados.

Uma outra medida tomada foi utilizar os dados processados dentro de um contexto coerente ao seu cenário de observação. Dados e opiniões não foram extrapolados e muito menos houve qualquer tipo de aglutinação e comparação de dados entre fontes diferentes. As inferências que foram atribuídas a um conjunto de trabalhos foram realizadas apenas na análise qualitativa e, ainda assim, considerava o conjunto geral dos trabalhos e não a comparação de dois deles.

Outra medida tomada para minimizar esta ameaça foi a utilização de o máximo de fontes possível para o mesmo evento. Vale ressaltar que para o *impeachment* isso não foi possível, já que não foram encontrados muitos estudos quantitativos a respeito do assunto. Mas, para as eleições de 2014 e 2018, nos dois turnos, houve uma série de trabalhos que traziam as mesmas indicações e informações complementares que corroboravam os resultados das pesquisas.

Outra medida para contextualizar os dados foi observar a evolução temporal dos fatos e dispor as pesquisas em uma espécie de “linha do tempo” para se construir uma narrativa dos fatos. Esse olhar contribuiu muito para a concepção da visão qualitativa e das evoluções entre cada episódio (eleições 2014 -> *impeachment* -> eleições 2018).

A segunda ameaça a validade identificada foi relacionada a **coleta dos dados** quantitativos trazidos ao bojo desse trabalho para compor o estudo de caso. Inicialmente o uso de engenhos de busca (*Google*) e bases de dados especializadas (*Google Acadêmico* e *LISA*) não garante a coleta de outras pesquisas que, potencialmente, poderiam enriquecer esta dissertação. Ao adotar este mecanismo, se permite que uma ameaça a validade externa ligada ao acesso à informação se faça presente nesta pesquisa.

Deve-se destacar a necessidade de cautela quanto a seleção das fontes, posto que os dados aqui apresentados são secundários, e foram gerados por institutos de pesquisas, entre eles Fundação Getúlio Vargas (FGV), Laboratório de estudos sobre Imagem e Cibercultura (Labic) da Universidade Federal do Espírito Santo (UFES) e imprensa, tais como o Jornal Estado de São Paulo (Estado) e Revista Exame, apesar de serem estudos especializados sobre o tema, não houve possibilidade de ser realizada a aferição direta dos valores e informações apresentados. Assim, foi desafiador definir quais dados seriam utilizados e de quais fontes seriam extraídos para reduzir ao máximo o viés da pesquisa, posto que as ações de desinformação verificada na Internet estão relacionadas as questões políticas e sociais que atendem a interesses de diversos atores de um contexto que demanda poder. Consequentemente, deve-se destacar que pesquisas encontradas em blogs e outros canais de credibilidade desconhecida na rede foram descartados, embora, não seja possível precisar a validade desta ação.

Assim, cabe repetir que a falta de acesso aos dados originais dificultou a análise do cenário da desinformação, sendo necessário um maior esforço para triangulação das informações que estavam fragmentadas, porém relacionadas a um mesmo contexto. Entretanto, no caso do *impeachment*, se fez necessário realizar uma coleta de dados manual, para complementar os dados realizados pelo estudo original. Todavia, esta mesma ação não foi realizada para 2014 e 2018, o que pode trazer algum tipo de desbalanceamento do que foi observado no *impeachment* e as outras eleições.

Isso pode ser observado como um agravante, posto que parte dos elementos que compõem essa pesquisa já não sejam tão facilmente encontrados, pois, tendem a se exaurir mais rapidamente nas plataformas da rede onde efetivamente foram produzidos em razão da velocidade com que se produz dados na Internet, fazendo com que aqueles que interessam para fins de coleta se percam no espaço virtual. Ou, porque as eleições de 2018 ainda eram recentes e estudos mais robustos foram publicados durante o processo de coleta.

Assim, quanto maior o lapso temporal entre o evento pesquisado nas fontes primárias e a coleta dos dados para a pesquisa, maior o desafio em obter esses dados de maneira consistente e completa. Desta forma, para a conclusão do estudo de caso foi necessário realizar inferências sobre o que foi coletado de modo a apresentar de

maneira lógica as conclusões trazidas nessa dissertação a partir das premissas que foram outrora destacadas por terceiros.

O que direciona ao último comentário dessa ameaça a validade que é a impossibilidade de verificar o rigor de pesquisa realizada pelos laboratórios e imprensa utilizados aqui como fontes primárias para este trabalho. Assim todas as ameaças a validade internas e externas dos estudos originais foram carregadas para a pesquisa.

Entretanto, ressalta-se que as ações realizadas para minimizar o impacto dos vieses, já apresentadas na ameaça a validade anterior, também auxiliam a minimizar as questões aqui apresentadas.

Por fim, a última ameaça a validade apresentada é quanto ao uso das **ferramentas botometer e twitter audit**, para verificar as contas do *Twitter*, se eram humanas ou automatizadas, que impactaram no evento do *impeachment*/ eleições 2018. Embora elas apontem uma indicação, não é possível afirmar, com 100% de certeza, se as informações que foram disseminadas partiram de perfis automatizados que nos limitou em alguns aspectos, tais quais:

(i) as ferramentas acima apenas sugerem se contas são automatizadas a partir de determinados comportamentos que, em tese, não são comumente realizados por humanos. Entretanto, não há a garantia de que tais perfis sejam mesmo automatizados, ou sejam ciborgues (semi-automatizados), ou mesmo 100% humano;

(ii) outro fator de impacto negativo no uso dessas ferramentas diz respeito ao fato dessas analisarem os perfis automatizados a partir das características atribuídas aos *bots* que, apesar de contar com traços que dizem respeito as máquinas sociais, não possuem todas as características dessas.

Apesar disso, o uso das referidas ferramentas demonstrou-se essencial, posto a necessidade de análise e identificação dos perfis atuantes nas campanhas de desinformação, e a ausência de outros meios que possibilitassem a realização da referida atividade.

6.4 Impactos Sociais

Conforme já apontado, a desinformação não é um acontecimento inédito na sociedade, entretanto nos moldes como hoje está configurada, contando com os recursos que corroboram sua ampla e veloz propagação, apresenta-se agora como um fenômeno antigo que se modernizou e ostenta contornos até então desconhecidos.

Nesse sentido, a pesquisa gera mudanças no que diz respeito ao conhecimento e discussão do fenômeno com maior propriedade, pois possibilita o reconhecimento desse como fator de impacto na sociedade, visto que seus desdobramentos atingem o cotidiano das pessoas até nas suas atividades diárias mais básicas.

Ou seja, ao entrar em contato com a pesquisa é possível observar e entender como agem as máquinas sociais no processo de desinformação, quais os resultados dessas ações e como evitar os prejuízos causados por elas. É a partir desse contexto que surge para o leitor a possibilidade de indagação na busca por posicionamentos claros sobre a política de informação e o uso de dados tanto das empresas administradoras das redes sociais, quanto do poder público, além de proporcionar o conhecimento de características das máquinas sociais enquanto agentes que atuam nas redes sociais, possibilitando a supervisão dos perfis automatizados que estão presentes na rede e a disseminação de conteúdo por esses veiculados.

Outro ponto que deve ser considerado é a conscientização social para que haja a utilização responsável dos meios de comunicação em rede, papel que deve ser desempenhado pelo cidadão também de maneira individual como meio de barrar a propagação de informações distorcidas, enviesadas ou falsas. Ressalta-se aqui que em grande parte a desinformação é impulsionada quando não há uma análise do que vai ser repassado pelo usuário (humano) nas redes sociais, sendo o encaminhamento de mensagens sem a devida filtragem fator que contribui para a perpetuação das informações de cunho duvidoso.

Percebe-se ainda que, em decorrência do que é apresentado nessa dissertação, a comunidade, acadêmica ou não, passará enxergar a manipulação ocasionada por ações de desinformação de maneira mais técnica, possibilitando a identificação desse fenômeno e garantindo poder de reação do receptor da mensagem, ao passo que poderá buscar informações para realizar a checagem voluntária dos fatos, bem como emitir alertas dentro da sua rede de contatos apontando que o conteúdo falta com a verdade e

consequentemente contendo as repercussões do fenômeno ainda que em pequena escala.

Quanto ao poder público, os resultados da pesquisa sugerem a necessidade de direcionar esforços e investimentos no campo da legislação, da pesquisa e tecnologia, vez que entre os principais impactos sociais causados por essas ações estão as investidas contra o Estado Democrático de Direito que afeta a ordem política, social e até econômica que deve ser precipuamente preservada pelo Estado.

Do ponto de vista legislativo, apesar de alguns avanços como a criação da Legislação Geral de Proteção de Dados (Lei nº 13.709/2018), é preciso mencionar as dificuldades de efetivação da própria lei, posto a ausência de instrumentos e meios que possibilitem a sua execução. É nesse sentido que surge a necessidade de pesquisas a partir dos resultados conhecido através desse trabalho, que possibilitará a criação de tecnologias (instrumentos) que auxiliem na busca por soluções ou permita ações que repreendam o mau uso das redes sociais para disseminação de conteúdo no sentido de desinformar.

Assim, considerando as descobertas trazidas como resultado dessa pesquisa, é possível estabelecer novos pontos de partidas a partir do que já foi construído, observando que, agora, o fenômeno da desinformação já se tornou conhecido quanto às suas características, manifestações e repercussões. Mais que isso, é possível distinguir quais ações que resultam em desinformação são na verdade praticadas por máquinas sociais (e não por humanos), tornando-se conhecida a capacidade desses agentes e seu modo de atuação.

Ou seja, o ponto de partida e descoberta desse fenômeno e suas variantes já estão postos, cabendo, a partir disso, seguir com a investigação e criação de instrumentos que sejam capazes de conter as repercussões negativas trazidas por esses agentes quanto utilizados para atender propósitos obscuros e ilegais.

6.5 Trabalhos Futuros

A temática trazida para o bojo dessa dissertação é extensa, razão pela qual diversos são os modos de abordagem através do qual poderia ser pesquisada. Nesse

toar, constata-se que as discussões aqui iniciadas não se esgotam, pelo contrário, teria uma infinidade de estudos de casos a serem abordados e outras proposições a serem formuladas através dos resultados obtidos com base nas observações de outros eventos de desinformação.

Isso se deve ao fato de que, à medida que outras campanhas de desinformação se propagam, é possível verificar uma especialização quanto ao *modus operandi* das empresas/atores que desfrutam/promovem esse tipo de serviço, seja pelo aperfeiçoamento das máquinas que atuam na proliferação de conteúdo, seja na forma de abordagem que se dá o contexto de disseminação das informações.

A exemplo disso, pode-se destacar que o estudo de caso trazido para discussão da presente pesquisa apresenta três eventos distintos que, apesar de tratar do contexto político brasileiro, é recepcionado e operado tanto por quem produz (*des*) informação quanto por quem recebe a informação (usuário) de formas diferentes. Observa-se nesse a migração de plataformas (*facebook -> twitter -> whatsapp*), bem como o desenvolvimento da expertise das máquinas e dos produtores de conteúdo, sugerindo-se que esse mercado vem atingindo uma progressiva evolução.

Desta forma, julga-se importante sugerir a **execução de outros estudos de casos** que permitam o aprofundamento das questões aqui levantadas, bem como o aperfeiçoamento dos termos e categorias criadas para tratar sobre desinformação. Ademais, observar outros contextos, sejam eles social, político ou econômico, trará a possibilidade de aplicar o entendimento estabelecido pela pesquisa ou até mesmo refutar pontos defendidos e destacados por essa.

Por fim, o monitoramento de outros eventos possibilitará a confirmação ou não daquilo que aqui se chama de padrão de comportamento das máquinas, como meio de identificar até que ponto sua capacidade de atuação nesse contexto vem sendo desenvolvida ou não; outra questão que deve ser somada a necessidade de novos estudos é detectar outras estratégias de desinformação que estão sendo aperfeiçoadas e aplicadas por quem produz esse tipo de conteúdo.

Sugere-se então como primeiro trabalho futuro a observação deste fenômeno em outros contextos que possibilite uma maior amplitude do leque da participação das máquinas no processo de desinformação.

A partir do que se foi observado nesta pesquisa, é possível notar a necessidade de **estabelecer ações preventivas** para a proteção do cidadão. Essas ações também devem estar previstas na seção dos trabalhos futuros, isso porque as discussões evidenciadas pela pesquisa apenas apontam a identificação de um problema, sua manifestação e repercussão no mundo prático, que demanda a necessidade de **novas pesquisas que apresentem soluções para efetivação de políticas de uso e disseminação da informação**.

Ou seja, é necessário alinhar os resultados obtidos as ações práticas que resultem na **criação de regulamentações** sobre questões que versem sobre a utilização de máquinas sociais para disseminação de informação e a criação de conteúdo com caráter inverídico. Atenta-se aqui para a necessidade de **criar mecanismos que permitam a execução da legislação** que, até certo ponto, proíbem os abusos ocasionados pela desinformação, como por exemplo o direito à informação e ao livre convencimento baseado na verdade, mas que no mundo prático pouco contribuem (atualmente) para conter a problemática.

Assim é sugerido como um segundo trabalho futuro a elaboração de regulamentações, políticas, mecanismos e legislações que tornem possível o controle das ações executadas por entidades não humanas ao mesmo tempo que proteja os cidadãos de abusos e proporcione uma maior segurança no uso da Internet.

REFEFÊNCIAS

ACKOFF, Russell L. Management misinformation systems. **Management science**, v.14, n. 4: p. 147-156. 1967. Disponível: <https://bit.ly/2ZqV8XT>. Acesso em: 22 jun. 2018.

ALLCOTT, Hunt; GENTZKOW, Matthew. Social media and fake news in the 2016 election. **Journal of economic perspectives**, v. 31, n. 2, p. 211-36, 2017. Disponível em: <https://bit.ly/2KYSnWh>. Acesso em: 22 jun. 2018.

ALVES, Rubem. **Entre a ciência e a sapiência**: o dilema da educação. São Paulo: Loyola, 2001.

AMERICAN LIBRARY ASSOCIATION. Resolution on Disinformation, Media Manipulation, & the Destruction of Public Information. 2005. Disponível em: <https://bit.ly/2WCN6cB>. Acesso em: 22 jun. 2018.

AMERICAN LIBRARY ASSOCIATION. Resolution on access to accurate information. 2017. Disponível em: <https://bit.ly/2lc8ol7>. Acesso em: 22 jun. 2018.

ANDERSON, Terry (Ed.). **The theory and practice of online learning**. Athabasca University Press, 2008.

ANDRADE, Graziela. **NÓS EM REDE**: informação, corpo e tecnologias. 2008. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

ANGELO, Edna. Redes sociais virtuais na sociedade da informação e do conhecimento: economia, poder e competência informacional. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 21, n. 46, 2016. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2016v21n46p71>. Acesso em: 26 jun. 2018.

AQUINO, Mirian de Albuquerque. A ciência da informação: novos rumos sociais para um pensar reconstrutivo no mundo contemporâneo. **Ci. Inf.** v. 36, n. 3, 2007. Disponível em: <https://bit.ly/2KNLNp5>. Acesso em: 22 jun. 2018.

ARIF, Ahmer. Designing to Support Reflection on Values & Practices to Address Online Disinformation. *In*: **Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing**. ACM, 2018. p. 61-64. Disponível em: <https://bit.ly/31uMj0S>. Acesso em: 22 jun. 2018.

ARNAUDO, Dan. Computational propaganda in Brazil: Social bots during elections. **Project on Computational Propaganda**, v. 8, 2017.

BAILLY, Christian. **Automata**: The Golden Age 1848-1914. London: Robert Hale, 2003.

BAKER, Vicky. Battle of the bots. **Index on Censorship**, v. 44, n. 2, p. 127-129, 2015. Disponível em: <https://bit.ly/2IICG6c>. Acesso em: 22 jun. 2018.

BALLARD, Susan. Cloud-watching robots: Douglas Bagnall's machine aesthetics. **Convergence**, v. 20, n. 3, p. 352-368, 2014.

BARTH, Bradley. **Malicious bot traffic climbs 9.5 percent in 2017**. 2018. Disponível em: <https://bit.ly/2Ggsnlx>. Acesso em: 11 jun. 2018

BBC. Na semana do impeachment, 3 das 5 notícias mais compartilhadas no Facebook são falsas. 2016. Disponível: <http://goo.gl/Vs8KRu>. Acesso em: 11 jun. 2018

BENKLER, Yochai; FARIS, Rob; ROBERTS, Hal; BOURASSA Nikki. **Understanding Media and Information Quality in an Age of Artificial Intelligence, Automation, Algorithms and Machine Learning**. 2018. Disponível em: <https://bit.ly/2RdzHXq>. Acesso em: 22 jun. 2018.

BERNERS-LEE, Tim; FISCHETTI, Mark. **Weaving the Web**: The original design and ultimate destiny of the World Wide Web by its inventor. DIANE Publishing Company, 2001.

BEZERRA, Arthur Coelho; CAPURRO, Rafael; SCHNEIDER, Marco. Regimes de verdade e poder: dos tempos modernos à era digital. **Liinc em Revista**, v. 13, n. 2, 2017. Disponível: <https://bit.ly/2CGurqH>. Acesso em: 05 fev. 2019.

BHARGAVA, Deepshikha. Intelligent agents and autonomous robots. *In*: **Rapid Automation: Concepts, Methodologies, Tools, and Applications**. IGI Global, 2019. p. 1134-1143.

BUCKLAND, Michael K. Information as thing. **Journal of the American Society for information science**, v. 42, n. 5, p. 351-360, 1991. Disponível em: <https://bit.ly/2XbW8BE>. Acesso em: 05 fev. 2019.

BURÉGIO, Vanilson; MEIRA, Silvio; ROSA, Nelson. Social machines: a unified paradigm to describe social web-oriented systems. *In*: **Proceedings of the 22nd International Conference on World Wide Web**. ACM, 2013. p. 885-890.

BURKHARDT, Joanna. **Social Media Bots**: How they spread misinformation. American Libraries, 2018. Disponível em: <https://bit.ly/2KI8IDk>. Acesso em: 09 jul. 2018.

CALAZANS, Angélica Toffano Seidel. Qualidade da informação: conceitos e aplicações. **TransInformação**, v. 20, n. 1, p. 29-45, 2008. Disponível em: <https://bit.ly/2lpsPae>. Acesso em: 22 jun. 2018.

CAPURRO, Rafael. What is information science for? A philosophical reflection. **Conceptions of library and information science**, p. 82-96, 1992.

CAPURRO, Rafael. **Foundations of Information Science: Review and Perspectives**. 2003. Disponível em: <https://bit.ly/2MO5mjH>. Acesso em: 9 jun. 2018.

CARLSEN, Steinar; GJERSVIK, Reidar. Organizational metaphors as lenses for analyzing workflow technology. **GROUP**, v. 97, p. 101-126, 1997.

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

CARVALHO, Olavo de. **O que é desinformação**. Rio de Janeiro: O globo. 17 de março de 2001. Disponível em: <https://olavodecarvalho.org/o-que-e-desinformacao/>. Acesso em: 9 jun. 2018.

CASTRO, César Augusto; RIBEIRO, Maria Solange Pereira. Sociedade da informação: dilema para o bibliotecário. **Transinformação**, v. 9, n. 1, p. 1725, janeiro/abril, 1997. Disponível em: <https://bit.ly/2N0ZrYL>. Acesso em: 9 jun. 2018.

CHAKRABORTY, Pinaki; SAXENA, Prem Chandra; KATTI, Chittaranjan Padmanabha. Fifty years of automata simulation: a review. **ACM Inroads**, v. 2, n. 4, p. 59-70, 2011.

CHAO, Chengcheng *et al.* **The spread of low-credibility content by social bots**. 2017. Disponível em: <https://arxiv.org/pdf/1707.07592.pdf>. Acesso em: 11 jun. 2018.

CHEONG, Pauline Hope; GONG, Jie. Cyber vigilantism, transmedia collective intelligence, and civic participation. **Chinese Journal of Communication**, v. 3, n. 4, p. 471-487, 2010. Disponível em: <https://bit.ly/2F4Cmxx>. Acesso em: 12 fev. 2019.

CHESNEY, Robert; CITRON, Danielle. Deepfakes e a nova guerra da desinformação: a nova era da geopolítica pós-verdade. **Aff.**, v. 98, 2019. Disponível em: <https://fam.ag/2R1nure>. Acesso em: 12 fev. 2019.

CHOI, Charles. **Automaton, Know Thyself: Robots Become Self-Aware**. 2011. Disponível em: <https://www.scientificamerican.com/article/automaton-robots-become-self-aware/>. Acesso em: 19 abr. 2019.

CHU, Zi *et al.* Detecting automation of twitter accounts: Are you a human, bot, or cyborg? **IEEE Transactions on Dependable and Secure Computing**, v. 9, n. 6, p. 811-824, 2012.

CORRÊA, Elisa Cristina Delfini; CUSTÓDIO, Marcela Gaspar. A informação enfurecida e a missão do bibliotecário em tempos de pós-verdade: uma releitura com base em Ortega y Gasset. **RBBB. Revista Brasileira de Biblioteconomia e Documentação**, v. 14, n. 2,

p. 197-214, 2018. Disponível em: <https://febab.emnuvens.com.br/rbbd/article/view/777>. Acesso em: 12 fev. 2019.

DEMO, Pedro. Ambivalências da sociedade da informação. **Ciência da Informação**, v. 29, n. 2, p. 37-42, 2000. Disponível em: <https://www.scielo.br/pdf/ci/v29n2/a05v29n2.pdf>. Acesso em: 12 mar. 2019.

DIAMOND, Larry. Liberation technology. **Journal of Democracy**, v. 21, n. 3, p. 69-83, 2010.

DONEDA, Danilo. Reflexões sobre proteção de dados pessoais em redes sociais. **Revista Internacional de Protección de Datos Personales**, n. 1, p. 1-12, 2012. Disponível: <https://bit.ly/2pPe4zK>. Acesso em: 09 jul. 2018.

DRETSKE, Fred I. Précis of Knowledge and the Flow of Information. **Behavioral and Brain Sciences**, v. 6, n. 1, p. 55-63, 1983. Disponível em: <https://bit.ly/2wQFMj0>. Acesso em: 12 mar. 2019.

DUNHAM, Ken; MELNICK, Jim. **Malicious bots**: an inside look into the cyber-criminal underground of the Internet. Auerbach Publications, 2008.

EDUARD, Laugaste. The Kalevela and the Kalevipoeg. *In*: Lauri Honko. **Religion, Myth and Folklore in the World's Epics**: The Kalevala and its Predecessors. Walter de Gruyter. pp. 265-286, 1990.

ENCARNAÇÃO, Bruno Pereira da. **A emergência das máquinas sociais na web, o caso futweet**. 2010. Dissertação (Mestrado em Ciência da Computação). – Centro de Informática, Universidade Federal de Pernambuco, Recife, 2010.

EXAME. 7 boatos da política brasileira que podem ter enganado você. 2015. Disponível em: <https://exame.com/brasil/7-boatos-da-politica-brasileira-que-podem-ter-enganado-voce/>. Acesso em: 23 ago. 2018.

FALLIS, Don. **A conceptual analysis of disinformation**. 2009. Disponível em: https://www.researchgate.net/publication/42101173_A_Conceptual_Analysis_of_Disinformation. Acesso em: 12 mar. 2019.

FALLIS, Don. What is disinformation? **Library Trends**, v. 63, n. 3, p. 401-426, 2015. Disponível em: <https://bit.ly/2XcuX9R>. Acesso em: 23 fev. 2018.

FERREIRA, Alexandre Valério; RIOS, José Riverson Araújo Cysne. Filtro bolha, câmara de eco e a formação de opiniões extremas. 2017. Disponível em: <https://portalintercom.org.br/anais/nacional2017/resumos/R12-0525-1.pdf>. Acesso: 25 ago. 2019.

FETZER, James H. Disinformation: The use of false information. **Minds and Machines**, v. 14, n. 2, p. 231-240, 2004. Disponível em: <https://bit.ly/2ZlvBzd>. Acesso em: 23 fev. 2018.

FGV. **Bots, Social networks and Politics in Brazil**: A study on illegitimate interferences with the public debate on the web, risks to the democracy and the 2018 elections. 2017. Disponível em: <http://goo.gl/GEfN1G>. Acesso: 29 set. 2018.

FGV. **Robôs, redes sociais e política no Brasil**: análise de interferências de perfis automatizados nas eleições de 2014. 2018a. Disponível em: <http://goo.gl/RTskE3>. Acesso: 29 set. 2018.

FGV. **O Estado da Desinformação**: Redes estrangeiras no twitter interferem no debate eleitoral brasileiro. 2018b. Disponível em: <http://goo.gl/7pCdDe>. Acesso: 29 set. 2018.

FISCHER, Steve. **Most Internet traffic comes from bots, not people**. 2017. Disponível em: <http://www.scielo.br/pdf/%0D/ci/v29n2/a05v29n2.pdf>. Acesso em: 11 jun. 2018.

FLETCHER, John. Deepfakes, Inteligência Artificial e Algum Tipo de Distopia: As Novas Faces do Desempenho Pós-Fato Online. **Theatre Journal**, v. 70, n. 4, p. 455-471, 2018.

FLINDERS, Matthew; WOOD, Matt. When politics fails: hyper-democracy and hyper-depoliticization. **New Political Science**, v. 37, n. 3, p. 363-381, 2015. Disponível em: <https://bit.ly/2Ki2Y2C>. Acesso em: 23 fev. 2018.

FLORIDI, Luciano. Brave. Net. World: The Internet as a disinformation superhighway? **The Electronic Library**, v. 14, n. 6, p. 509-514, 1996. Disponível em: <https://bit.ly/2IHM8qp>. Acesso em: 23 fev. 2018.

FLORIDI, Luciano. Concepções semânticas de informação. **Stanford Encyclopedia of Philosophy**. 2005. Disponível em: <https://stanford.io/2IHG4OC>. Acesso em: 08 fev. 2019.

FLORIDI, Luciano. **The philosophy of information**. New York: Oxford University Press. 2011. Disponível em: <https://bit.ly/2X1UX7S>. Acesso em: 08 fev. 2019.

FLORIDI, Luciano. Steps forward in the philosophy of information. **Etica & Política**, 14(1), p. 304-310. 2012. Disponível em: <https://bit.ly/31sTuql>. Acesso em: 08 de jan. 2019.

FLORIDI, Luciano. **The fourth revolution**: how the infosphere is reshaping human reality. Oxford, UK: Oxford University Press, 2014.

FLORIDI, Luciano. **The onlife manifesto**: Being human in a hyperconnected era. Springer Nature, 2015.

FLORIDI, Luciano. Artificial intelligence, deepfakes and a future of ectypes. **Philosophy & Technology**, v. 31, n. 3, p. 317-321, 2018. Disponível em: <https://bit.ly/2X50BWF>. Acesso em: 08 fev. 2019.

FLORIDI, Luciano; ILLARI, Phyllis. **The philosophy of information quality**. Cham, CH: Springer, 2014.

FOX, Christopher J. *Information and Misinformation*. Westport, Connecticut: Greenwood Press, 1983.

FREITAS, Carlos S.; BENEVENUTO, Fabrício; VELOSO, Adriano. Socialbots: Implicações na segurança e na credibilidade de serviços baseados no Twitter. *In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS (SBRC)*, 32, 2014, Florianópolis. Proceedings... SBC: Porto Alegre, 2014. Disponível em: <https://bit.ly/2NuWqfk>. Acesso em: 05 jul. 2018.

FREND, Steven J.; NICHOLS, Rebecca M.; LOFTUS, Elizabeth F. Current issues and advances in misinformation research. **Current Directions in Psychological Science**, v. 20, n. 1, p. 20-23, 2011. Disponível em: <https://bit.ly/2ldU0Rt>. Acesso em: 12 jun. 2018.

GEMROT, Jakub; BROM, Cyril; PLCH, Tomáš. A periphery of pogamut: From bots to agents and back again. *In: International Workshop on Agents for Games and Simulations*. Springer, Berlin, Heidelberg, 2010. p. 19-37.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GOH, Dion H.; LEE, Chei Sian. Perceptions, quality and motivational needs in image tagging human computation games. **Journal of Information Science**, v. 37, n. 5, p. 515-531, 2011.

GORWA, Robert; GUILBEAULT, Douglas. Unpacking the social media bot: A typology to guide research and policy. **Policy & Internet**, 2018.

GOTTFRIED, Jeffrey; SHEARER, Elisa. **News Use Across Social Media Platforms 2016**. Pew Research Center, 2016.

HADDOW, Gill *et al.* Cyborgs in the everyday: Masculinity and biosensing prostate cancer. **Science as Culture**, v. 24, n. 4, p. 484-506, 2015.

HARAWAY, Donna. Manifesto Ciborgue. *In: SILVA, Tomaz Tadeu (Org.). Antropologia do Ciborgue: as vertigens do pós-humano*. Belo Horizonte: Autêntica, 2009. Disponível em: <https://www.passeidireto.com/arquivo/24858047/haraway-donna-manifesto-ciborgue>. Acesso em: 27 fev. 2019.

HASELTON, Martie G.; NETTLE, Daniel; ANDREWS, P. W. The evolution of cognitive bias. *In*: BUSS, David M. (Ed.). **The Handbook of Evolutionary Psychology**. Hoboken, NJ, US: John Wiley & Sons Inc, 2005.

HAUSTEIN, Stefanie *et al.* Tweets as impact indicators: Examining the implications of automated “bot” accounts on Twitter. **Journal of the Association for Information Science and Technology**, v. 67, n. 1, p. 232-238, 2016. Disponível em <https://bit.ly/2lkjsBI>. Acesso em: 23 jun. 2018.

HENDLER, James; MULVEHILL, Alice M. **Social machines**: the coming collision of artificial intelligence, social networking, and humanity. Apress, 2016.

HERNON, Peter. Disinformation and misinformation through the Internet: Findings of an exploratory study. **Government Information Quarterly**, v. 12, n. 2, p. 133-139, 1995. Disponível em: <https://bit.ly/2F8qcE3>. Acesso em: 13 maio 2018.

HIGGINS, Kathleen. Post-truth: a guide for the perplexed. **Nature News**, v. 540, n. 7631, 2016. Disponível em: <https://go.nature.com/2OGcque>. Acesso em: 24 out. 2018.

HILBERT, Martin. Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. **Psychological bulletin**, v. 138, n. 2, 2012. Disponível em: <https://bit.ly/31x95VW>. Acesso em: 24 out. 2018.

HAO, Karen. **Congress wants to protect you from biased algorithms, deepfakes, and other bad AI**. MIT Technology Review. 2019. Disponível em: <https://bit.ly/2UY6dRK>. Acesso em: 13 mar. 2019.

HUPSEL FILHO, Valmar; GALHARDO, Ricardo. Governo cita uso de robôs nas redes sociais em campanha eleitoral. Estadão de São Paulo. 2015. Disponível em: <http://goo.gl/RkxFcR>. Acesso: 15 set. 2018.

IBGE. Instituto Brasileiro de Geografia e Estatística (IBGE). **Pesquisa nacional por amostra de domicílios**: síntese de indicadores 2015 / IBGE, Coordenação de Trabalho e Rendimento. Rio de Janeiro. 2016. Disponível em: <https://bit.ly/2rqiYEB>. Acesso em: 03 mai. 2018.

ISHIGURO, Hiroshi. Android science. Cognitive Science Society, Osaka, 2005.

ITS. **Poder Computacional**: Automação no uso do WhatsApp nas Eleições. 2019. Disponível: <http://goo.gl/S7Ei78>. Acesso em: 29 jan. 2019.

ITU. About International Telecommunication Union. **Relatório da ICT Facts & Figures 2016**. Disponível em: <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2016.pdf>. Acesso em: 03 mai. 2018.

JOHNSTON, Les. What is vigilantism? **The British Journal of Criminology**, v. 36, n. 2, p. 220-236, 1996. Disponível em: <https://academic.oup.com/bjc/article-abstract/36/2/220/563555>. Acesso em: 24 out. 2018.

PIEROZZI JÚNIOR, Ivo *et al.* Análise de dinâmica de uso e de desempenho: o caso do web site da Embrapa Monitoramento por Satélite. **Ci. Inf**, v. 32, n. 1, p. 102-114, 2003. Disponível em: <https://www.scielo.br/pdf/ci/v32n1/15978.pdf>. Acesso em: 19 jan. 2019.

LUPA. Só 4 das 50 imagens mais compartilhadas por 347 grupos de WhatsApp são verdadeiras. 2018. Disponível em: <http://goo.gl/Qy622p>. Acesso em: 19 jan. 2019.

KARLOVA, Natascha A.; FISHER, Karen E. "Plz RT": A social diffusion model of misinformation and disinformation for understanding human information behaviour. **Information Research**, v. 18, n. 1, p. 1-17, 2013. Disponível em: <https://bit.ly/2HwbloX>. Acesso em: 24 out. 2018.

KIETZMANN, Jan *et al.* Social media? Get serious! Understanding the functional building blocks of social media. **Business horizons**, v. 54, n. 3, p. 241-251, 2011. Disponível em: <https://bit.ly/2HvDwlf>. Acesso em: 24 out. 2018.

LABIC. Laboratório de Estudos sobre Imagem e Cibercultura. **Bots contra a sociedade**. 2014. Disponível em: <https://bit.ly/2OpZzBg>. Acesso em: 12 out. 2018.

LAGRANDEUR, Kevin. **Androids and intelligent networks in early modern literature and culture**: Artificial slaves. Reino Unido: Routledge, 2012.

LAMPE, Cliff; ELLISON, Nicole; STEINFELD, Charles. A Face (book) in the crowd: Social searching vs. social browsing. *In: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM, 2006. p. 167-170. Disponível em: <https://bit.ly/2XlubOR>. Acesso em: 03 mai. 2018.

LARA, Gabriel Teixeira *et al.* IMPLICAÇÕES POLÍTICAS, SOCIAIS E ÉTICAS DA POPULARIZAÇÃO DE DEEPFAKES POR MEIO DE SOFTWARES LIVRES. *In: Anais do Congresso Nacional Universidade*, EAD e Software Livre. v. 2, n. 9, 2018. p. 1-6. Disponível em: <https://bit.ly/2WynPeN>. Acesso em: 24 out. 2018.

LAZER, David *et al.* The science of fake news. **Science**, v. 359, n. 6380, 2018. p. 1094-1096. Disponível em: <https://bit.ly/31lbRO9>. Acesso em: 20 jan. 2019.

LEBOW, David G. A SOCIAL MACHINE FOR TRANSDISCIPLINARY RESEARCH. **Informing Science**, v. 21, 2018. Disponível em <http://www.inform.nu/Articles/Vol21/ISJv21p201-217Lebow4185.pdf>. Acesso em: 16 abr. 2019.

LEITE, Leonardo Ripoll Tavares; MATOS, José Claudio. Zumbificação da informação: a desinformação e o caos informacional. *In: Anais do Congresso Brasileiro de*

Biblioteconomia, Documentação e Ciência da Informação-FEBAB. 2017. Disponível em: <https://bit.ly/2tyN0Gs>. Acesso em: 20 mai. 2018.

LEVY, Pierre. **Inteligência coletiva**. Edições Loyola, 2007.

LEWANDOWSKY, Stephan *et al.* Misinformation and its correction: Continued influence and successful debiasing. **Psychological Science in the Public Interest**, v. 13, n. 3, p. 106-131. 2012. Disponível em: <https://bit.ly/2WD9H8P>. Acesso em: 17 set. 2018.

LOFTUS, Elizabeth F. Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. **Learning & Memory**, v. 12, n. 4, p. 361-366, 2005. Disponível em: <https://bit.ly/31vchS7>. Acesso em: 17 set. 2018.

MAANI, Kambiz. Systems Thinking and Internet. pp 2719-2723. *In*: KOSROW, Mehdi (Ed.). **Encyclopedia of Information Science and Technology** v. 1, 2005.

MACHADO, Caio *et al.* News and political information consumption in Brazil: Mapping the first round of the 2018 Brazilian presidential election on Twitter. **Computational Propaganda Project**, 2018.

MAGRANI, Eduardo. **Democracia conectada**: a Internet como ferramenta de engajamento político-democrático. Curitiba: Juruá, 2014. 222p. Disponível em: <https://bit.ly/2RMDc74>. Acesso em: 17 set. 2018.

MAHMOUD, Qusay H. **Software Agents**: Characteristics and Classification. School of Computing Science, Simon Fraser University, p. 1-12, 2000.

MALINI, Fabio. **17 de abril**: quando a Câmara dos Deputados virou alvo do #impeachmentday. 2016. Disponível em: <https://medium.com/@fabiomalini/17-de-abril-quando-a-c%C3%A2mara-dos-deputados-virou-alvo-a2a6f0e2f48b>. Acesso em: 25 ago. 2019.

MARKINES, Benjamin; CATTUTO, Ciro; MENCZER, Filippo. Social spam detection. *In*: **Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web**. ACM, 2009. p. 41-48. Disponível em: <https://bit.ly/2WFht1W>. Acesso em: 17 set. 2018.

MARSHALL, Jonathan Paul *et al.* **Disorder and the Disinformation Society**: The Social Dynamics of Information, Networks and *Software*. Routledge, 2015.

MASCARO, Olivier; SPERBER, Dan. The moral, epistemic, and mindreading components of children's vigilance towards deception. **Cognition**, v. 112, n. 3, p. 367-380, 2009. Disponível em: <https://bit.ly/2WN6fh6>. Acesso em: 11 dez. 2018.

MATHEUS, Renato Fabiano. Rafael Capurro e a filosofia da informação: abordagens, conceitos e metodologias de pesquisa para a Ciência da Informação. **Perspect. ciênc. inf.**, Belo Horizonte, v. 10, n. 2, p. 140-165, jul./dez. 2005.

MCLURE, Helen. The wild, wild web: The mythic American West and the electronic frontier. **Western Historical Quarterly**, v. 31, n. 4, p. 457-476, 2000. Disponível em: <https://bit.ly/2WCzPAW>. Acesso em: 12 set. 2018.

MEIRA, Silvio R. L. *et al.* The Emerging Web of Social Machines. In: **Proceedings of the 2012 IEEE 35 th Annual Computer Softwares and Applications Conference**. 2011. Disponível em: <https://ieeexplore.ieee.org/document/6032321>. Acesso em: 12 set. 2019.

MERCHANT, Arpit; JHA, Tushant; SINGH, Navjyoti. The use of trust in social machines. In: **Proceedings of the 25th International Conference Companion on World Wide Web**. International World Wide Web Conferences Steering Committee, 2016.

MICHEL, Maria Helena. **Metodologia e Pesquisa Científica em ciências sociais**: um guia prático para acompanhamento da disciplina e elaboração de trabalhos monográficos. São Paulo: Atlas, 2009.

MOSSERI, Adam. News feed fyi: Addressing hoaxes and fake news. **Facebook newsroom**, v. 15, p. 12, 2016.

NASCIMENTO, Leandro Marques do *et al.* A new architecture description language for social machines. In: **Proceedings of the 23rd International Conference on World Wide Web**. ACM, 2014. p. 873-874.

NELSON, Jacob L.; TANEJA, Harsh. The small, disloyal fake news audience: The role of audience availability in fake news consumption. **New media & society**, v. 20, n. 10, p. 3720-3737, 2018. Disponível em: <https://bit.ly/2XGIbcV>. Acesso em: 12 set. 2018.

NING, Huansheng *et al.* Cyberlogic paves the way from cyber philosophy to cyber science. **IEEE Internet of Things Journal**, v. 4, n. 3, p. 783-790, 2017.

NÓBREGA, Lívia. Representação robótica na Ficção Científica: implicações subjetivas. In: **Anais Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação**, 2016, p. 1-16.

NUNES, Amanda Maria de Almeida *et al.* A AÇÃO DOS BOTS NO PROCESSO DE DESINFORMAÇÃO EM ELEIÇÕES E REFERENDOS. **XIX Encontro Nacional de Pesquisa em Ciência da Informação**, Marília, 2018. Disponível em: http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1069. Acesso em: 13 jan. 2019.

NWANA, Hyacinth S. *Software agents: An overview*. **The knowledge engineering review**, v. 11, n. 3, p. 205-244, 1996.

OLIVEIRA, Carina Capelão de; VIEIRA, Carolina Coimbra; LOPES, Ananda de Oliveira. Critérios para seguir usuários e bots no Twitter. **Múltiplos Olhares em Ciência da Informação-ISSN 2237-6658**, v. 5, n. 2, 2016.

OLIVEIRA, Éric Tadeu Camacho de *et al.* The Influence of Retweeting Robots During Brazilian Protests. *In: 2016 49th Hawaii International Conference on System Sciences (HICSS)*. IEEE, 2016. p. 2068-2076.

OLIVEIRA, Janderson Rodrigo de. **Um sistema integrado para navegação autônoma de robôs móveis**. 2010. 102 f. Dissertação (Mestrado) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2010.

PALERMOS, Spyridon Orestis. Social machines: a philosophical engineering. **Phenomenology and the Cognitive Sciences**, v. 16, n. 5, p. 953-978, 2017. Disponível em: <https://link.springer.com/content/pdf/10.1007%2Fs11097-016-9489-4.pdf>. Acesso em: 10 mar. 2019.

PARISER, Eli. **The filter bubble**: What the Internet is hiding from you. Penguin UK, 2011. Disponível: <https://bit.ly/2XHLbow>. Acesso em: 13 fev. 2019.

PARK, Enno. Ethical issues in cyborg technology: Diversity and inclusion. **NanoEthics**, v. 8, n. 3, p. 303-306, 2014.

PINHEIRO, Marta Macedo Kerr; BRITO, Vladimir de P. Em busca do significado da desinformação. **DataGramaZero Revista de Informação**, v. 15, n. 6, 2014. Disponível em: <https://www.brapci.inf.br/index.php/res/v/8068>. Acesso em: 20 abr. 2018.

PORTERFIELD, Jason. **Robots, Cyborgs, and Androids**. The Rosen Publishing Group, Inc, 2018.

PRODANOV, Cleber Cristiano. **Metodologia do trabalho científico**: métodos e técnicas da pesquisa e do trabalho acadêmico. Novo Hamburgo: Feevale, 2013.

RAUNIG, Gerald. **Bodies, Things, and Social Machines**. 2008. Disponível em: <https://transversal.at/pdf/journal-text/228/>. Acesso em: 20 abr. 2018

RECUERO, Raquel. Redes sociais na Internet: considerações iniciais. *In: Encontro dos núcleos de pesquisa da XXVII Intercom*, 4, Porto Alegre, RS, 2004. **Anais...** Porto Alegre: INTERCOM, 2004.

RÉGIS, Fátima. Os Autômatos da Ficção Científica: reconfigurações da tecnociência e do imaginário tecnológico. **Intexto**, n. 15, p. 41-56, 2006.

ROGERS, Todd *et al.* Artful paltering: The risks and rewards of using truthful statements to mislead others. **Journal of Personality and Social Psychology**, v. 112, n. 3, p. 456, 2017. Disponível em: <https://bit.ly/2XKaQNa>. Acesso em: 20 abr. 2018.

ROURE, David *et al.* Observing Social Machines Part 2: How to Observe? *In: Proceedings of the ACM Web Science Conference*. ACM, 2015. Disponível: <https://dl.acm.org/citation.cfm?id=2786475>. Acesso em: 18 fev. 2019.

ROUSH, Wade. **Social Machines**: Computing means connecting. 2005. Disponível em: <https://www.technologyreview.com/s/404466/social-machines/>. Acesso em 26 mar. 2019.

SANTAELLA, Lucia. Culturas e artes do pós-humano: da cultura das mídias à cibercultura. São Paulo: Paulus, 2004. Disponível em: https://drive.google.com/file/d/0ByGOj9_gW1Y7QUptSGxkOFI3dDDQ/view. Acesso em: 27 fev. 2019.

SANTANA, Célio Andrade Júnior; LIMA, Camila Oliveira. O papel das máquinas sociais na formação de opinião em rede| The Role of Social Machines in Opinion Formation on Web. **Liinc em Revista**, v. 13, n. 2, 2018. Disponível em: <https://bit.ly/2XKUVOL>. Acesso em: 28 jun. 2018.

SANTANA, Célio *et al.* O papel da ciência da informação para o desenvolvimento do conceito de máquinas sociais. *In: VI Encontro Ibérico EDICIC 2013*. 2013. Disponível em: <http://ocs.letras.up.pt/index.php/EDICIC/edicic2013/paper/viewPaper/407>. Acesso em: 25 fev. 2019.

SANTOS, Eduardo Ferreira dos. Rumo a “social machines” sobre dados criminais. 2015. Dissertação (Mestrado Profissional em Computação Aplicada) – Universidade de Brasília, 2015. Disponível em: <https://core.ac.uk/download/pdf/33551796.pdf>. Acesso em: 16 abr. 2019.

SANTOS, Stéfanie *et al.* Do século XVI para o XXII: Shakespeare em mangá. **Revista Língua&Literatura**, v. 19, n. 34, p. 71-83, 2017.

SARACEVIC, Tefko. **Quality of information services**: Considerations for library & information services in the networked world. 1999. 35 slides. Disponível em: <https://bit.ly/2trBvkU>. Acesso em: 21 jun. 2018.

SCARANTINO, Andrea; PICCININI, Gualtiero. Information without truth. **Metaphilosophy**, v. 41, n. 3, p. 313-330, 2010. Disponível em: <https://bit.ly/2XQJWDH>. Acesso em: 21 jun. 2018.

SCHRADER, Alvin. M. The domain of information science: problems in conceptualization and in consensus-building. **Information Services & Use**, v. 6, n. 5-6, p. 169-205, 1986. Disponível em: <https://bit.ly/2XKzu0a>. Acesso em: 21 jun. 2018.

SEMMEHACK, Peter. **Social machines**: how to develop connected products that change customers' lives. John Wiley & Sons, 2013.

SHADBOLT, Nigel. Knowledge acquisition and the rise of social machines. **International Journal of Human-Computer Studies**, v. 71, n. 2, p. 200-205, 2013. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1071581912001607>. Acesso em: 10 mar. 2019.

SHADBOLT, Nigel *et al.* The rise of social machines: the development of a human/digital ecosystem. **IEEE Consumer Electronics Magazine**, v. 5, n. 2, p. 106-111, 2016. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7450779>. Acesso em: 17 fev. 2019.

SHIRKY, Clay. The political power of social media: Technology, the public sphere, and political change. **Foreign affairs**, p. 28-41, 2011. Disponível em: <https://bit.ly/2XI3Osj>. Acesso em: 21 jun. 2018.

SHU, Kai *et al.* Fake news detection on social media: A data mining perspective. **ACM SIGKDD Explorations Newsletter**, v. 19, n. 1, p. 22-36, 2017. Disponível em: <https://bit.ly/2TC3Ccg>. Acesso em: 6 set. 2018.

SILVA FILHO, Rubens da Costa; SILVA, Leila Morás; LUCE, Bruno. Impacto da pós-verdade em fontes de informação para a saúde. **RBB. Revista Brasileira de Biblioteconomia e Documentação**, v. 13, p. 271-287, 2017. Disponível em: <https://bit.ly/2LSRwLf>. Acesso em: 6 set. 2018.

SILVA, Nayane Maria Rodrigues. Fake News: a revitalização do jornal e os efeitos Fact-Checking e CrossCheck no noticiário digital. **Temática**, v. 13, n. 8, p. 16-19, 2017. Disponível em: <https://bit.ly/31xtuuc>. Acesso em: 6 set. 2018.

SILVERMAN, Craig; SINGER-VINE, Jeremy. Most Americans who see fake news believe it; new survey says. **BuzzFeed News**, 2016. Disponível em: <https://bit.ly/2KixWrs>. Acesso em: 6 set. 2018.

SKYRMS, Brian. **Signals**. New York: Oxford University Press. 2010. 208p.

SMART, Paul R.; SHADBOLT, Nigel R. Social machines. *In*: **Encyclopedia of Information Science and Technology, Third Edition**. IGI Global, 2015. p. 6855-6862.

STABLEFORD, Brian. **Science fact and science fiction**: an encyclopedia. CRC Press. pp. 22-23. 2006.

STAHL, Bernd Carsten. On the Difference or Equality of Information, Misinformation, and Disinformation: A Critical Research Perspective. **Informing Science**, v. 9, 2006.

STARBIRD, Kate *et al.* Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *In: IConference 2014 Proceedings*, 2014. Disponível em: <https://bit.ly/2XKf5s4>. Acesso em: 6 set. 2018.

THEOI. **The automatones of Greek Mythology**. 2000. Disponível em: <https://www.theoi.com/Ther/Automotones.html>. Acesso em: 6 set. 2018.

THURLER, Larriza. **BOTS E GESTÃO DO CONHECIMENTO: UMA BREVE ANÁLISE DO CENÁRIO ATUAL E TENDÊNCIAS**. P. 1-8, 2015. Disponível em: <https://bit.ly/2tMEy6z>. Acesso em: 28 jun. 2018.

TOLOSA, Gabriel; BORDIGNON, Fernando. **Revisión: tecnologia de agentes de software**. *Ciência da Informação*, v. 28, n. 3, 1999.

TROTTIER, Daniel (Ed.). Social media, politics and the state: Protests, revolutions, riots, crime and policing in the age of facebook, twitter and youtube. *In: In Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in an Age of Facebook, Twitter and YouTube*. Routledge, p. 189-206, 2014.

TUDJMAN, Miroslav; MIKELIC, Nives. Information science: Science about information, misinformation and disinformation. **Proceedings of Informing Science+ Information Technology Education**, p. 1513-1527, 2003.

TURING, Alan M. **Can a machine think**. *Mind*, 1950.

VAN DIJCK, José *et al.* **The Platform Society: Public Values in a Connective World**. 2018.

VOLKOFF, Vladimir. **Pequena história da desinformação: do cavalo de Tróia à Internet**. Curitiba: Vila do Príncipe, 2004.

WARDLE, Claire. The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. **Digital Journalism**, v. 6, n. 8, p. 951-963, 2018. Disponível em: <https://bit.ly/2MLyAj9>. Acesso em: 6 fev. 2019.

WATSON, Bruce *et al.* **United States Intelligence: An Encyclopedia**. New York/London: Garland Publishing, Inc. 1990.

WEHMHOENER, Karl Allen. Social norm or social harm: An exploratory study of Internet vigilantism. 2010. Disponível em: <https://bit.ly/31vjhhQ>. Acesso em: 6 fev. 2019.

ZATTAR, Marianna. Competência em informação e desinformação: critérios de avaliação do conteúdo das fontes de informação. **Liinc em Revista**, v. 13, n. 2, 2017. Disponível em: <https://bit.ly/2Xhf4PG>. Acesso em: 11 jun. 2018.