UMA PROPOSTA PARA IMPULSIONAR AS PESQUISAS EM GOVERNANÇA DE ALGORITMOS DE IA E DADOS

Julião Braga[*]
Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, SP, BR
juliao.braga@ufabc.edu.br

Francisco Regateiro[†]
Instituto Superior Técnico
Universidade de Lisboa
Lisboa, PT
francisco.regateiro@tecnico.ulisboa.pt

Itana Stiubiener[‡]
Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, SP, BR
Itana.stiubiener@ufabc.edu.br

Juliana Cristina Braga[§]
Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, SP, BR
juliana.braga@ufabc.edu.br

ABSTRACT

This paper is a proposal to research on governance of artificial intelligence algorithms and data from a comparative study of relevant works and also from experience in Internet governance. Complementarily, the results from this study will be made available for discussion and improvement in a decentralized control environment, implemented especially for this purpose.

1 Introdução e justificativa, com síntese da bibliografia fundamental

Inteligência Artificial (IA) é o termo usado para sistemas computacionais que tentam imitar aspectos da inteligência humana, incluindo funções que associamos intuitivamente à inteligência, tais como aprendizagem, resolução de problemas, pensamento e ação racional [1]. De um modo geral e independente da aplicação, estes sistemas são considerados uma caixa preta resultando em informações assimétricas entre os seus desenvolvedores e seus consumidores [2]. Um dos exemplos mais tristes e que evidenciam a consequência desta assimetria é o projeto do sistema MCAS[5] do Boing 737 MAX, que levou a dois acidentes com 346 mortes em outubro de 2018 (Lion Air) e março de 2019 (Ethiopian Airlines). Quando o ângulo do sensor de ataque falhou, os algoritmos embutidos forçaram o avião a baixar o nariz, resistindo às repetidas tentativas dos pilotos, confusos, de virar o nariz para cima.

Ben Shneiderman, em seu livro *Human-Centered AI*, que comenta os dois acidentes com o Boing 737 MAX, considera que o futuro destes algoritmos de IA é centrado no ser humano, principalmente tornando-se super ferramentas, que amplificam as habilidades humanas, capacitando as pessoas de forma notável mas, ao mesmo tempo, garantindo o controle humano [3]. Ben nomeou estes algoritmos com a sigla **HCAI**, acrônimo do título de seu livro.

Há inúmeras outras aplicações usando IA, como por exemplo, aquelas de habitam a Internet, que se comportam de forma desproporcional. Uma descrição detalhada dos chamados vieses algorítmicos está no livro de Safya Noble, Algoritmos da Opressão [4].

Informação assimétrica, vieses e outras questões estão incomodando os desenvolvedores, pesquisadores e a outros interessados, todos determinados a descobrir o que está faltando [5]! Perspectivas associadas com ética [6, 7, 8, 9],

[*] http://lattes.cnpq.br/7092085044582071
[†] https://fenix.tecnico.ulisboa.pt/homepage/ist13522
[‡] http://lattes.cnpq.br/4008970012663480
[§] http://lattes.cnpq.br/7111526592323456
[5] Acrônimo de *Manoeuvring Characteristics Augmentation System*

regulamentação [10, 11, 12, 13, 14], governança [15, 16, 17, 18, 2, 19] e muitas outras [20, 21, 22, 23, 24, 25, 26, 27, 28] estão na pauta de todas as partes interessadas, em busca de alternativas apropriadas – por exemplo, estas questões estão amplamente debatidas em [3].

O objetivo desta proposta é estudar o estado da arte em algoritmo e governança de dados, produzindo um documento que, além de agregar os insights mais significativos, reconhece e compara os resultados da experiência da governança da Internet, um domínio de aplicação tradicional, mas determinante. Acreditamos que esta comparação elucidará as premissas para uma governança comprovadamente benéfica [29].

Em paralelo, pretendemos divulgar e disponibilizar as conclusões propostas em um ambiente adequado para o debate das partes interessadas e consequente

A governança da Internet e o ambiente de debates, são discutidos nas sub-seção 1.1; O conceito de Organização Autônoma Decentralizada (DAO) é apresentado na seção 1.2. Na sequência, a sub-seção 1.3 apresenta os estudos primários e secundários que servirão como base para a bibliografia selecionada, alinhados com os objetivos do presente trabalho.

1.1 A Internet

A Internet é uma rede de redes de computadores. Tais redes de computadores são denominadas de Sistemas Autônomos (ASes) e possuem uma identificação única. Os ASes estão interligados por meio de uma complexa infraestrutura de telecomunicações, espalhada por toda a terra e além. A operacionalização dos ASes que usam esta infraestrutura de telecomunicações é feita por programas, isto é, algoritmos denominados genericamente de protocolos.

Portanto, a Internet é uma organização decentralizada, isto é, não há controle central e sua estrutura pode ser vista, abstratamente, na Figura 1.

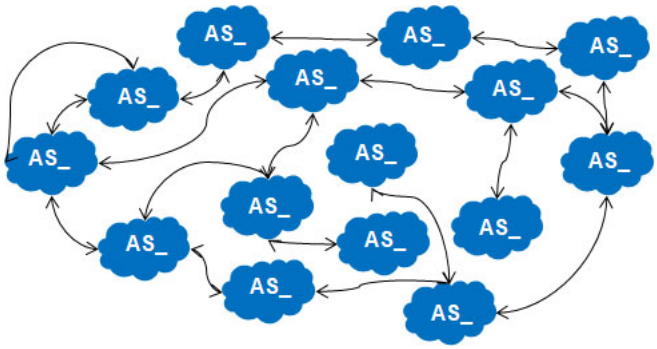

Figure 1: Como a Internet se organiza. Fonte: [30]

O seu complexo ambiente de governança e as instituições nele envolvidas são exibidos na Figura 2.

A necessidade de se organizar para distribuir consistentemente os sistemas de numeração únicos (números de ASes, números IPv4 e IPv6, entre outros [31, 32]), definidos em seus documentos de padronização (as chamadas RFCs[6]) foi a principal motivação para se estabelecer o modelo de governança [33].

Por outro lado, os milhares de voluntários interessados em produzir algoritmos para a infraestrutura da Internet, se reúnem através de grupos de correiro eletrônico e presencialmente, três vezes por ano, em torno de uma instituição denominada *Internet Engineering Task Force*[7] (IETF) e fazem-no motivados pelo objetivo de tornar a Internet cada vez melhor. Os algoritmos são definidos em documentos preliminares e submetidos a um consenso entre os voluntários distribuídos em grupos que aprovam ou não a proposta apresentada. É como se fossem artigos avaliado por muitos pares, de forma mais rigorosa do que os tradicionais dois ou três pares avaliando artigos tradicionais! Os documentos, aprovados ou ainda em processo de aprovação – estes sendo alterados em vários momentos – são submetidos e preservados em um repositório controlado por um grupo reunido sob o nome de RFC Editor[8], e é parte do processo de governança dos protocolos da Internet, como se pode ver no retângulo tracejado da Figura 2.

[6] Acrônimo de *Request for Comments*
[7] `https://www.ietf.org/`
[8] Organizado em `https://www.rfc-editor.org/`

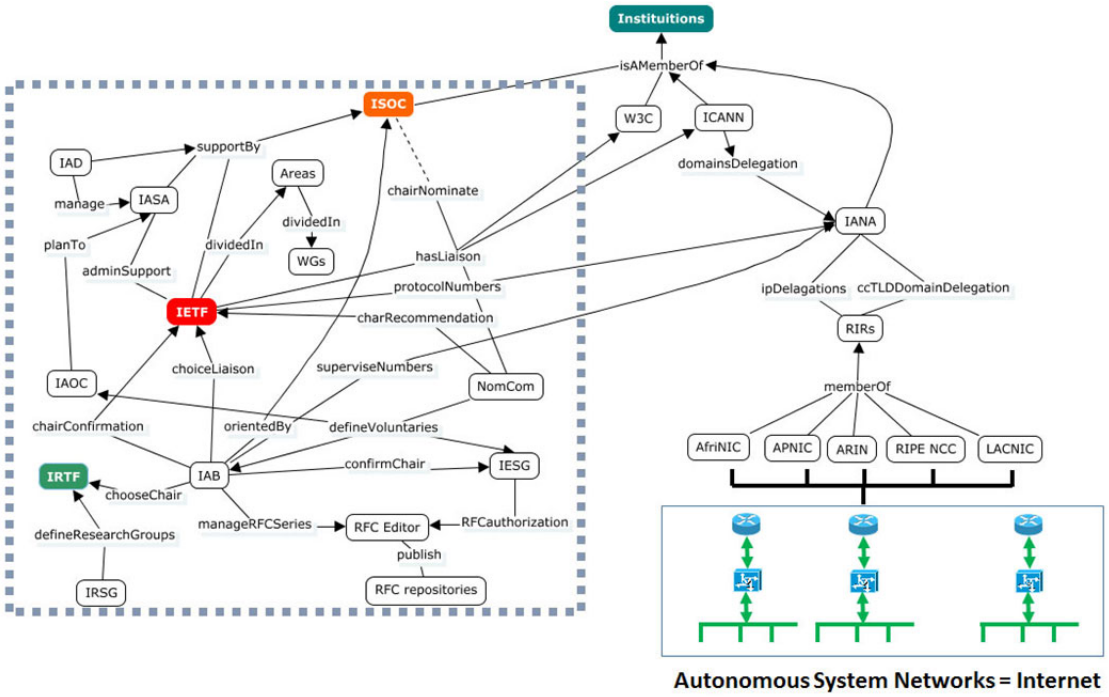

Figure 2: Ecosistema da governança da Internet. Fonte: [30]

No desenvolvimento do projeto, será estudado o comportamento das instituições e pessoas que fazem parte do ecossistema de governança da Internet, a fim de entender como elas são capazes de garantir o desenvolvimento e a manutenção dos algoritmos subjacentes ao conjunto dos chamados protocolos. Além disso, o estudo incluirá uma caracterização comparativa dos pontos relevantes para a governança encontrados nos algoritmos das aplicações que tenham qualquer técnica de IA anexada.

No desenvolvimento do projeto será estudado o comportamento das instituições e pessoas que fazem parte do ecossistema de governança da Internet, a fim de entender como elas são capazes de garantir/proteger o desenvolvimento e manutenção dos algoritmos subjacentes ao conjunto dos chamados protocolos. Neste aspecto, o estudo deve expressar um caráter comparativo, capturando os pontos relevantes para a governança dos algoritmos - com possíveis técnicas de IA em anexo.

1.2 *Decentralized Autonomous Organization* (DAO)

Em outubro de 2008, Satoshi Nakamoto, um pseudônimo, publicou um artigo intitulado *Bitcoin: A Peer-to-Peer Electronic Cash System*, onde previa a necessidade de estabelecimento de um sistema de pagamento na Internet, com características peer-to-peer e que seja autônomo, seguro e limitado em quantidade. Surgiu, então, a Bitcoin lançada em 3 de janeiro de 2009, onde cada Bitcoin custava US$0,0008 [34]. Hoje, quase quatro anos depois, um Bitcoin custa US$21,276.36 e existem 20.930 cryptomoedas diferentes[9].

Nakamoto propôs que a Bitcoin fosse implementada sobre uma estrutura de dados chamada *blockchain*. Uma *blockchain* ou cadeia de blocos é um conjunto de estruturas de dados do tipo lista, cujos itens (blocos) são encadeados ao contrário, formando uma base de dados distribuída, com o conteúdo dos blocos e suas indicações de encadeamento, criptografadas, tornando esta estrutura encadeada inviolável e indestrutível. Esta cadeia de blocos, na realidade, é como se fosse um livro razão contábil e público, de todas as transações ou eventos digitais que foram executados e compartilhados entre as partes interessadas. Cada transação do livro razão é verificada por consenso de uma maioria de seus participantes, que produziram blocos no sistema encadeado de dados [35].

Com o passar dos anos, a experiência da cadeia de blocos do Bitcoin evoluiu e novas técnicas de pesquisa e manipulação das cadeias apareceram sendo que uma nova moeda surgiu, juntamente com a inclusão de programas (isto é, algoritmos) dentro dos blocos, passíveis de serem executados. Esta evolução se deu com o surgimento do *Ethereum*,

[9]https://coinmarketcap.com/

uma moeda que fortaleceu a cadeia de blocos como uma estrutura completamente decentralizada e garantindo a anonimidade de seus usuários.

E a esta nova propriedade de decentralização com garantia do anonimato (além de inviolável e indestrutível) deu-se o nome de *Decentralized Finance* (DeFi). Algumas vezes DeFi confunde-se com a referência a **web3**. A ela se juntaram-se uma grande quantidade de aplicações inovadoras, entre elas, em 2016, a Organização Autônoma Decentralizada[10,11] (DAO), com estrutura exibida na Figura 3

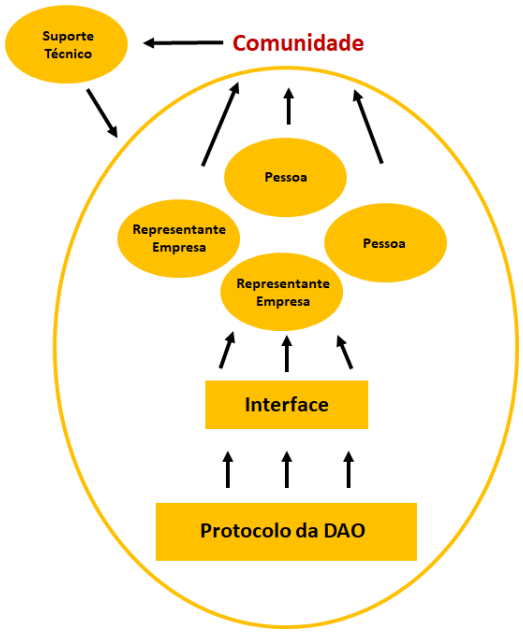

Figure 3: Estrutura proposta para a DAO. Adaptação da apresentação de Moriah Pacheco Rickli, da Ribon, no Lift Learning[12]

Para se construir uma DAO é necessário um planejamento (isto é, um *design*) muito bem elaborado e isto será feito dando origem a um documento detalhando tal planejamento e ao qual se dará o nome de *Human-Algorithm agreement*, onde as regras estabelecidas em contratos inteligentes (algoritmos processáveis na web3) serão caracterizadas. O algoritmos, residentes nos blocos da cadeia sobre a qual a DAO será implementada, dispensa a necessidades de humanos controladores. Este documento e o código compatível com as regras serão escritos conforme especificados no cronograma (Tabela 2) e ocorrerão nos primeiros meses do projeto, com o apoio das partes interessadas reunidas através do *Discord*[13]. O ambiente do *Discord*, denominado Humano-Algorithm pode ser acessado através da URL convite `https://discord.gg/sKG8ypS7ed`.

A Figura 4, mostra um conjunto de DAOs construídas até 2021, caracterizadas por áreas de interesse. Em particular, a DAO **Uniswap**[14] é um ecossistema de usuários, desenvolvedores, *designers* e educadores, muito parecida com a DAO que está se propondo neste trabalho. Hayden Adams e colegas descrevem a última versão do protocolo da Uniswap [36].

1.3 Bibliografia Fundamental

Na Tabela 1 está referenciada a literatura a ser analisada para que se possa entender o mecanismo de governança de algoritmos e dados e, permitir que haja um comparativo, também, das propostas recomendadas.

As referências estão classificadas em oito categorias e não se esgotam na relação apresentada nesta proposta:

[10] Tradução livre do inglês de: *Decentralized Autonomous Organization*
[11] `https://bit.ly/daoland`
[12] `https://ribon.io` e `https://bit.ly/lift-06`
[13] O Discord é um ambiente de colaboração *on-line*, muito utilizado pelos usuários da web3.
[14] `https://uniswap.org/community`

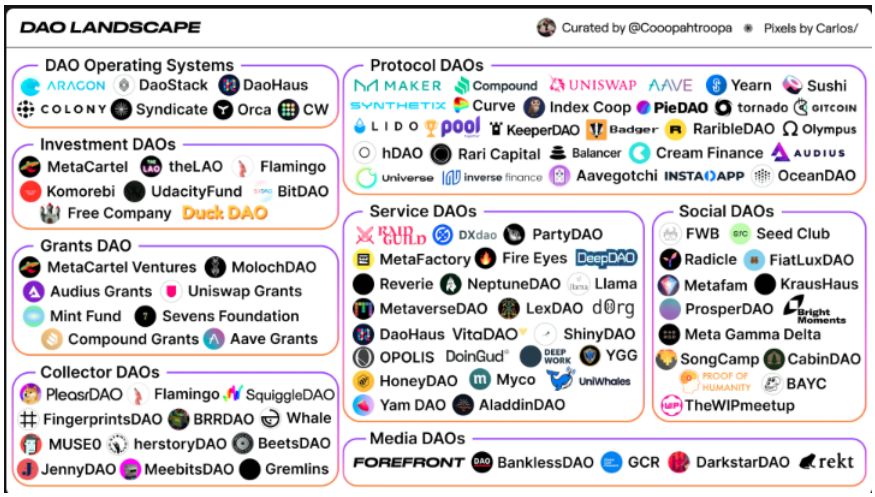

Figure 4: Em 2021 existiam mais de 100 DAOs em operação, manipulando acima de US$10 bilhões em ativos. Fonte: `https://bit.ly/daoland`

Table 1: Estudos primários e secundários de interesse do projeto, a serem avaliados

#	Referências	Classificação
1.	[37],[38],[39],[33],[40],[41],[42],[43],[44],[45], [46],[47]	Internet
2.	[48],[49],[50],[51],[52],[53],[54],[55],[56],[57],[58],[59],[60],[61],[62],[63]	Algoritmos
3.	[64],[65],[66],[67],[68],[69],[70],[71],[72],[73],[74],[75],[76],[77],[78]	
4.	[79],[80],[81],[82].[83],[84],[85],[86],[87],[88],[89],[90],[91],[92],[93],[94]	
5.	[95],[96],[97],[98],[99],[100],[101],[102],[103],[104],[105],[106],[107],[108]	
6.	[109][110],[111],[112],[113],[114],[115],[116],[117],[118],[119],[120],[121]	
7.	[122][123],[124],[125],[126],[127],[128],[129],[130],[131],[132],[133],[134]	
8.	[135],[136],[137],[138],[139],[140],[141],[142],[143],[144],[145],[146],[147]	
9.	[148],[149],[150],[151],[152],[153],[154],[155],[156],[157],[158],[159],[160]	
10.	[161],[162],[163],[164],[165],[166],[167],[168],[169],[170],[171]	
11.	[172],[173],[174],[175],[176],[177],[178],[179],[180],[181],[182],[183]	DAO
12.	[184],[185],[186],[187],[188],[189],[190],[191],[192],[193],[194],[195],[196]	Economia
13.	[197],[198],[199],[200],[201],[202],[203],[204],[205],[206],[207],[208],[209]	
14.	[210]	
15.	[211],[212],[213],[214],[215],[216],[217],[218],[219],[220],[221],[222],[223]	Outros
16.	[224],[225],[226],[227],[228],[229],[230],[231],[232],[233],[234],[235],[236]	
17.	[237],[238],[239],[240],[241],[242],[243],[244],[245],[246],[247],[248],[249]	
18.	[250],[251],[252],[253],[254],[255],[256],[257],[258],[259],[260],[261],[262]	
19.	[263],[264],[265],[266],[267],[268]	
20.	[269],[11],[270],[271],[272],[273],[274],[275],[276],[277],[278],[279],[280]	Governança
21.	[281],[282],[283],[284],[285],[286],[287],[288],[289],[290],[291],[292],[293]	
22.	[294],[295],[296],[297],[298],[299],[300],[301],[302],[303],[304],[305],[306]	
23.	[307],[308],[309],[310],[311],[312],[313],[314],[315],[316],[317],[318],[319]	
24.	[320],[321],[322],[323],[324],[325],[326],[327],[328],[329],[330],[331],[332]	
25.	[333],[334],[335],[336],[337],[338],[339],[340],[341],[342],[343],[344],[345]	
26.	[346],[347],[348],[349],[350],[351],[352],[353],[354],[355],[356],[357],[358]	
27.	[359],[360],[361],[362],[363],[364]	
28.	[365],[366],[367],[368]	RLiteratura
29.	[369],[370],[371],[372],[373],[374],[375],[376],[20],[377],[378],[379],[380]	Social
30.	[381],[382],[383],[384],[385],[386],[387],[388],[389],[390]	

1. **Internet:** Incluem as referências que abordam o tema de governança da Internet.
2. **Algoritmos:** São as referências que exibem algoritmos de IA em diversas áreas de aplicação.

3. **DAO:** Referências que abordam as DAOs e respectivas técnicas sobre as quais elas são construídas (blockchain e criptomoedas).
4. **Economia:** Referências que abordam questões relacionadas com a economias dos algoritmos e seus ambientes.
5. **Outros:** Um conjunto de referências que descrevem o envolvimento de algoritmos de IA aos assuntos: Bots, Discriminação, Engenharia de Software, IA, Jogos, Robótica e Segurança.
6. **RLiteratura:** Trabalhos de revisão de literatura, incluindo revisões sistemáticas.
7. **Social:** Textos que referenciam os aspectos social, ético e filosóficos dos algoritmos.

Na complementação da descrição do projeto, a seção 2 apresenta o Plano de Trabalho com o respectivo cronograma. A seção 3 descreve a metodologia de desenvolvimento do projeto. A seção 4 apresenta as alternativas para análise dos resultados. A seção 5, expressa os agradecimentos dos autores. O projeto termina com a apresentação completa das referências bibliográficas adicionais consolidadas até a dada de disponibilização deste artigo.

2 Plano de Trabalho e Cronograma

Na tabela 2 as principais etapas do projeto são discriminadas, com seus respectivos envolvimentos nos meses de duração do projeto, cujo marco inicial será a primeira semana de outubro de 2022.

Table 2: Cronograma do Projeto

#	Discriminação	Datas								
		10/22	**11/22**	**12/22**	**01/23**	**02/23**	**03/23**	**04/23**	**05/23**	**6/23**
1.	Descrição do projeto	●								
2.	Desenvolvimento do artigo	●	●	●	●	●	●	●		
3.	Submissão do artigo								●	
4.	Avaliar publicação em *preprint*								●	
5.	Seleção da Literatura	●	●	●	●					
6.	Avaliação da Literatura	●	●	●	●	●	●	●		
7.	Implementação do Discord	●								
8.	Avaliação de equipe/*startup*	●	●	●						
9.	*Human-Algorithm agreement*	●	●	●	●	●	●			
10.	Desenvolvimento da DAO	●	●	●	●	●	●	●	●	●
11.	Implementação da DAO				●					
12.	Teste da DAO				●	●	●	●	●	●
13.	Recursos *off-chain* da DAO	●	●	●	●	●	●	●	●	●
14.	Disponibilização da DAO					●				
15.	Inclusão de novas referências	●	●	●	●	●	●	●		

Há uma versão em inglês deste artigo em `https://doi.org/10.31219/osf.io/sr7kt` [391].

3 Metodologia

O levantamento bibliográfico do presente trabalho foi obtido a partir das referências distribuídas pelo Catedra Oscar Sala. Este levantamento pode crescer no decorrer do projeto, com base em outras referências obtidas a partir de levantamentos não-sistemáticos. A avaliação será qualitativa e comparativa sobre as referências que proponham diretamente governança de algoritmo de IA e/ou de dados. Isto indica que a análise do texto se preocupará com a subjetividade do pesquisador e destacará as diferenças e similaridades entre as propostas. Nessas referências que não tocam em questões de governança, a preocupação se voltará para o aspecto relacionado à contribuição que ela faz para a proposta de governança.

Durante o desenvolvimento do trabalho proposto neste texto, serão coletadas contribuições da colaboração entre as partes interessadas, que é apoiada por dois ambientes: (1) o repositório na estrutura da Open Science Foundation (OSF) [392] e, (2) o grupo Discord, cujo convite para participar é `https://discord.gg/nMUdgeUT`.

O gerenciamento do projeto será feito através do ProjectLibre[15].

[15] https://www.projectlibre.com/

A implementação da DAO será incumbência dos autores, com eventual apoio de equipe especializada ou uma *startup*, ambas em fase de avaliação. A implementação deve garantir o comportamento da DAO aderente aos seus *smart contracts* e as especificações descritas no *agreement* referenciado acima.

Durante a implementação do DAO, haverá um esforço de programação usando a linguagem *Solidity*[16] para adaptar os *smart contracts* originais [393]. Esta adaptação será feita analisando os *smart contracts* contra a proposta do DAO caracterizada pelo *agreement*.

O processo de avaliação do DAO implementado, com as modificações dos contratos inteligentes, será feito através das partes interessadas, na própria DAO, no *Discord* e/ou no repositório público disponível no OSF [392].

4 Análise dos resultados

O acompanhamento do projeto será feito:

- Pelos orientadores do projeto, no desenvolvimento do artigo no Overleaf,
- Através do uso da implementação da DAO,
- Através do ambiente publico na OSF [392],
- Pelos Orientadores, no ambiente privativo de colaboração da equipe de desenvolvimento do projeto, criado na OSF [394] e,
- Pelo arquivo do ProjectLibre, cujo cronograma será publicado no ambiente privativo na OSF e atualizado semanalmente.

5 Agradecimentos

Os autores agradecem ao Professor Dr. Virgílio de Almeida, catedrático da Cátedra Oscar Sala, do Instituto de Estudos Avançados da USP, pela viabilidade do projeto. Também agradecem ao Carlos Benedito Lima de Menezes Júnior, da Ribon, pelo apoio na consolidação das ideias em torno da DAO.

References

[1] Sanjit A Seshia, Dorsa Sadigh, and S Shankar Sastry. Toward verified artificial intelligence. *Communications of the ACM*, 65(7):46–55, 2022.

[2] Urs Gasser and Virgilio A.F. Almeida. A Layered Model for AI Governance. *IEEE Internet Computing*, 21(6):58–62, 2017.

[3] Ben Shneiderman. *Human-Centered AI*. Oxford University Press, 2022.

[4] Safya. Noble and Felipe Damorim. *Algoritmos da Opressão: Como os mecanismos de busca reforçam o racismo*. Editora Rua do Sabão, Rio de Janeiro, 1 edition, 2022.

[5] M Mitchell Waldrop. What are the limits of deep learning? *Proceedings of the National Academy of Sciences*, 116(4):1074–1077, 2019.

[6] Katina Michael, Diana Bowman, Meg Leta Jones, and Ramona Pringle. Robots and socio-ethical implications [guest editorial]. *IEEE Technology and Society Magazine*, 37(1):19–21, 2018.

[7] Andrea Censi, Konstantin Slutsky, Tichakorn Wongpiromsarn, Dmitry Yershov, Scott Pendleton, James Fu, and Emilio Frazzoli. Liability, ethics, and culture-aware behavior specification using rulebooks. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8536–8542. IEEE, 2019.

[8] Michael Kearns and Aaron Roth. *The ethical algorithm: The science of socially aware algorithm design*. Oxford University Press, 2019.

[9] Flávio S Corrêa da Silva and Nina S. T. Hirata. Inteligência Ética. *Computação Brasil*, 7:15–18, 2022.
After presenting the two complementary approaches, present in AI activities - the symbolic, based on logical systems, and the adaptive, based on machine learning - they discuss what an ethical system is, as well as norms and values for leading a worthwhile life, which originates in Ancient Greece, and define the three different classes of these systems: based on virtues, on duties, and on the consequence of actions.

[16] `https://en.wikipedia.org/wiki/Solidity`

They then relate ethics to the development of AI systems. They also discuss the need for formulating technical requirements and ways to measure how intelligent systems meet these requirements. Finally, they discuss some initiatives in this direction.

[10] Christianr Mölle and Arnaud Amouroux, editors. *Governing the Internet: Freedom and Regulation in the OSCE Region*. Organization for Security and Co-operation in Europe (OSCE), 2007.

[11] Eduardo Bismarck. PL 21/2020, 2020.

[12] Europe Commission. Coordinated Plan on Artificial Intelligence 2021 Review. Technical report, European Commission, Brussels, 2021.

[13] Margrethe Vestager and Thierry Breton. Uma Europa Preparada para a Era Digital : Comissão propõe novas regras e ações para promover a excelência e a confiança na inteligência artificial. Technical report, Comissão Euroeia, 2021.

[14] Sociedade Brasileira de Computação. Ética e regulação na inteligência artificial. Technical report, Sociedade Brasileira de Computação, 7 2022.

[15] Solon Barocas, Sophie Hood, and Malte Ziewitz. Governing Algorithms: A Provocation Piece. *SSRN Electronic Journal*, pages 1–12, 2013.

[16] Florian Saurwein, Natascha Just, and Michael Latzer. Governance of algorithms: Options and limitations. *Info*, 17(6):35–49, 2015.

[17] Danilo Doneda and Virgilio A.F. Almeida. What Is Algorithm Governance? *IEEE Internet Computing*, 20(4):60–63, 2016.

[18] Lucas D. Introna. Algorithms, Governance, and Governmentality: On Governing Academic Writing. *Science Technology and Human Values*, 41(1):17–49, 2016.

[19] Martin Ebers and Marta Cantero Gamito, editors. *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges*. Springer, 2021.

[20] Adam D.I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 111, pages 8788–8790, 2014.

[21] David Lazer. The rise of the social algorithm. *Science*, 348(6239):1090–1091, 2015.

[22] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

[23] Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan. The social dilemma of autonomous vehicles. *Science*, 352(6293):1573–1576, 2016.

[24] Matthew O Jackson. *The human network: How your social position determines your power, beliefs, and behaviors*. Pantheon Books, 2019.

[25] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. The science of fake news. *Science*, 359(6380):1094–1096, 2018.

[26] European Data Protection Supervisor (EDPS). Towards a new digital ethics. Technical Report September, European Organization, 2015.

[27] Anne Magaly de Paula Canuto. Ética no Uso de Dados Biométricos: Histeria ou uma Preocupação Coerente? *Computação Brasil*, 7:36–39, 2022.

[28] Mengyi Wei and Zhixuan Zhou. Ai ethics issues in real world: Evidence from ai incident database. *arXiv preprint arXiv:2206.07635*, 2022.

[29] Stuart Russell. Provably beneficial artificial intelligence. In *27th International Conference on Intelligent User Interfaces*, pages 3–3, 2022.

[30] Juliao Braga, Jeferson Campos Nobre, Lisandro Zambenedetti Granville, and Marcelo Santos. Como Protocolos Inovadores são Criados e Adotados em Escala Mundial: Uma visão sobre o Internet Engineering Task Force (IETF) e a Infraestrutura da Internet. In Taisy Silva Weber and Claudia Aparecida Martins, editors, *Jornadas de Atualiza cão em Informática 2020*, page 45. Sociedade Brasileira de Computa cão, Cuiabá, MT Brazil, 2020. Available in: https://doi.org/10.5753/sbc.5728.3.2.

[31] ICANN. Internet Corporation for Assigned Names and Numbers. `http://www.icann.org`. Acessado em 12/09/2013.

[32] IANA. Internet Assigned Numbers Authority, 2014. `http://www.iana.org`. Acessado em 03/05/2020.

[33] Diego Rafael Canabarro and Flavio Rech. A Governança da Internet: Definição, Desafios e Perspectivas. In *9o ENCONTRO DA ABCP*, page 17, 2014.

[34] Satoshi Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. *Decentralized Business Review*, page 9, 2008.

[35] Michael Crosby, Nachiappan, Pradan Pattanayak, Sanjeev Verma, and Vignesh Kalyanaraman. BlockChain Technology: Beyond Bitcoin. *Applied Innovation Review*, 2016.

[36] Hayden Adams, Noah Zinsmeister, Moody Salem moody, Uniswaporg River Keefer, and Dan Robinson. Uniswap v3 Core. Technical Report March, Uniswap, 2021.

[37] Janet Abbate. *Inventing the Internet*. MIT Press, 1999.

[38] Lucas Andrade, Juliao Braga, Stefany Pereira, Rafael Roque, and Marcelo Santos. In-Person and Remote Participation Review at IETF. In *Proceeding of CSBC 2018 - V Workshop pre IETF*, page 11, Natal, RN Brazil, July 2018. To be published. Available at: http://braga.net.br/papers/In-Person%20and%20Remote%20Participation%20Review%20at%20IETF.pdf.

[39] Lee A. Bygrave and Jon Bing. *Internet Governance: Infrastructure and Institutions*. Oxford University Press, New York, 2009.

[40] Diego Rafael Canabarro. *Governança global da internet: tecnologia, poder e desenvolvimento*. Doutoral thesis, Federal University of Rio Grande do Sul, 2014.

[41] Alexandre Arns Gonzales. *Quem Governa a Governança da Internet? Uma análise do papel da Internet sobre os rumos do sitema-mundo*. Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, 2016.

[42] Fabrício Pasquot Bertini Polido and Lucas Costa Dos Anjos, editors. *Marco Civil E Governança Da Internet: Diálogos Entre O Doméstico E O Global*. Faculdade de Direito da UFMG, 2016.

[43] Fabrício Bertini Pasquot Polido, Lucas Cosda dos Anjos, and Luíza Couto Chaves Brandão, editors. *Tecnologias e Conectividade: Direito e Políticas na Governança das Redes*. IRIS, 2017.

[44] Wolfgang Kleinwächter. The History of Internet Governance. In *Governing the Internet: Freedom and Regulation in the OSCE Region*, pages 41—-65. OSCE Region, Vienna, 2007.

[45] Christian Moller. Governing the Domain Name System: An Introduction to Internet Infrastructure. In *Governing the Internet: Freedom and Regulation in the OSCE Region*, pages 29–39. Organization for Security and Co-operation in Europe (OSCE), Vienna, 2007.

[46] Alexandre Pacheco da Silva, Ana Paula Camelo, Diego R. Canabarro, and Flavio Rech Wagner, editors. *Estrutura e funcionamento da internet : aspectos técnicos, políticos e regulatórios*. FGV, 2021.

[47] Elizabeth Machado Veloso. Legislação sobre Internet no Brasil. Technical report, Camara dos Deputados, 2009.

[48] Martin Adam, Michael Wessel, and Alexander Benlian. Ai-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2):427–445, 2021.

[49] Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and mobile computing*, 7(6):643–659, 2011.

[50] Saar Alon Barkat and Madalina Busuioc. Human-ai interactions in public sector decision-making: Automation bias and selective adherence to algorithmic advice. *Accepted Manuscript*, 2022.

[51] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–13, 2019.

[52] Saleema Amershi, Andrew Begel, Christian Bird, Robert DeLine, Harald Gall, Ece Kamar, Nachiappan Nagappan, Besmira Nushi, and Thomas Zimmermann. Software engineering for machine learning: A case study. In *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, pages 291–300. IEEE, 2019.

[53] Tawfiq Ammari, Jofish Kaye, Janice Y Tsai, and Frank Bentley. Music, search, and iot: How people (really) use voice assistants. *ACM Trans. Comput. Hum. Interact.*, 26(3):17–1, 2019.

[54] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. Concrete problems in ai safety. *arXiv preprint arXiv:1606.06565*, 2016.

[55] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine Bias, 2016.

[56] George A. Akerlof and Rachel E. Kranton. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press, Princeton, 1 edition, 2010.

[57] Kenneth C Arnold, Krysta Chauncey, and Krzysztof Z Gajos. Predictive text encourages predictable writing. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, pages 128–138, 2020.

[58] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. The moral machine experiment. *Nature*, 563(7729):59–64, 2018.

[59] Lisanne Bainbridge. Ironies of automation. In *Analysis, design and evaluation of man–machine systems*, pages 129–135. Elsevier, 1983.

[60] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 'it's reducing a human being to a percentage' perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems*, pages 1–14, 2018.

[61] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems*, 29, 2016.

[62] Josh Bongard, Victor Zykov, and Hod Lipson. Resilient machines through continuous self-modeling. *Science*, 314(5802):1118–1121, 2006.

[63] N Bolstrom. *Superintelligence. Paths, dangers, strategies*. Oxford University Press, United Kingdom, 2014.

[64] Jeffrey M Bradshaw, Robert R Hoffman, David D Woods, and Matthew Johnson. The seven deadly myths of "autonomous systems". *IEEE Intelligent Systems*, 28(3):54–61, 2013.

[65] Elizabeth Broadbent. Interactions with robots: The truths we reveal about ourselves. *Annual review of psychology*, 68(1):627–652, 2017.

[66] Connor Brooks and Daniel Szafir. Visualization of intended assistance for acceptance of shared control. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11425–11430. IEEE, 2020.

[67] Connor Brooks and Daniel Szafir. Balanced information gathering and goal-oriented actions in shared autonomy. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 85–94. IEEE, 2019.

[68] Miles Brundage, Shahar Avin, Jasmine Wang, Haydn Belfield, et al. Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. *arXiv preprint*, 2020.

[69] Erik Brynjolfsson and Tom Mitchell. What can machine learning do? Workforce implications. *Science*, 358(6370):1530–1534, 2017.

[70] Joy Buolamwini and Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*, pages 77–91. PMLR, 2018.

[71] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, 2017.

[72] Rafael A Calvo, Dorian Peters, Karina Vold, and Richard M Ryan. Supporting human autonomy in ai systems: A framework for ethical enquiry. In *Ethics of Digital Well-Being*, pages 31–54. Springer, 2020.

[73] Neil A. H. Campbell. The Evolution of Flight Data Analysis. In *Proceedings of Australian Society of Air Safety Investigators*, pages 1–22, 2003.

[74] Juliano Cappi. *Internet, Big Data e discurso de ódio: reflexões sobre as dinâmicas de interação no Twitter e os novos ambientes de debate político*. Doctoral thesis, Pontifícia Universidade Católica de São Paulo, 2017.

[75] Felix Carros, Johanna Meurer, Diana Löffler, David Unbehaun, Sarah Matthies, Inga Koch, Rainer Wieching, Dave Randall, Marc Hassenzahl, and Volker Wulf. Exploring human-robot interaction with the elderly: results from a ten-week case study in a care home. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[76] Radhika Chemuturi, Farshid Amirabdollahian, and Kerstin Dautenhahn. Adaptive training algorithm for robot-assisted upper-arm rehabilitation, applicable to individualised and therapeutic human-robot interaction. *Journal of NeuroEngineering and Rehabilitation*, 10(1), 2013.

[77] Bo Chen, Chunsheng Hua, Bo Dai, Yuqing He, and Jianda Han. Online control programming algorithm for human–robot interaction system with a novel real-time human gesture recognition method. *International Journal of Advanced Robotic Systems*, 16(4):1–18, 2019.

[78] Lu Cheng, Kush R Varshney, and Huan Liu. Socially responsible ai algorithms: Issues, purposes, and challenges. *Journal of Artificial Intelligence Research*, 71:1137–1181, 2021.

[79] Alexandra Chouldechova, Diana Benavides-Prado, Oleksandr Fialko, and Rhema Vaithianathan. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In *Conference on Fairness, Accountability and Transparency*, pages 134–148. PMLR, 2018.

[80] Paul F. Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, pages 4300–4308, 2017.

[81] Jacob W Crandall, Mayada Oudah, Fatimah Ishowo-Oloko, Sherief Abdallah, Jean-François Bonnefon, Manuel Cebrian, Azim Shariff, Michael A Goodrich, Iyad Rahwan, et al. Cooperating with machines. *Nature communications*, 9(1):1–12, 2018.
textbfThis study examines algorithmic cooperation with humans and provides an example of methods that can be used to study the behaviour of human–machine hybrid systems.

[82] Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503–507, 2015.

[83] Nilesh Dalvi, Pedro Domingos, Sumit Sanghai, and Deepak Verma. Adversarial classification. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 99–108, 2004.

[84] Rajarshi Das, James E Hanson, Jeffrey O Kephart, and Gerald Tesauro. Agent-human interactions in the continuous double auction. In *International Joint Conference on Artificial Intelligence*, volume 17, pages 1169–1178. Lawrence Erlbaum Associates Ltd, 2001.

[85] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240, 2006.

[86] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114, 2015.

[87] Julia Dressel and Hany Farid. The accuracy, fairness, and limits of predicting recidivism. *Science advances*, 4(1):1–5, 2018.

[88] Douglas C Engelbart and William K English. A research center for augmenting human intellect. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 395–410, 1968.

[89] Douglas C Engelbart. Augmenting human intellect: A conceptual framework. *Menlo Park, CA*, page 21, 1962. Reprinted in Packer, R. amd Kprdam. L;. eds; (2–1). Multimedia: From Wagner to Virtual Reality. New York: W. W. Norton, 64-90.

[90] Danielle Ensign, Sorelle A Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. Runaway feedback loops in predictive policing. In *Proceedings of Machine Learning Research*, pages 1–12, 2018.

[91] Ziv Epstein, Blakeley H Payne, Judy Hanwen Shen, Abhimanyu Dubey, Bjarke Felbo, Matthew Groh, Nick Obradovich, Manuel Cebrian, and Iyad Rahwan. Closing the ai knowledge gap. *arXiv preprint arXiv:1803.07233*, 2018.

[92] Michael Feldman, Sorelle A Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. Certifying and removing disparate impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 259–268, 2015.

[93] Sophie Freeman, Martin Gibbs, and Bjørn Nansen. 'don't mess with my algorithm': Exploring the relationship between listeners and automated curation and recommendation on music streaming services. *First Monday*, 2022.

[94] Tarleton Gillespie. The Relevance of Algorithms. In *Media technologies: Essays on communication, materiality, and society*, pages 167–194. The MIT Press, 2014.

[95] Enric Galceran, Alexander G Cunningham, Ryan M Eustice, and Edwin Olson. Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction: Theory and experiment. *Autonomous Robots*, 41:1367–1382, 2017.

[96] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. Datasheets for datasets. *Communications of the ACM*, 64:86–92, 2021.

[97] Alessandro Giusti, Jérôme Guzzi, Dan C Cireşan, Fang-Lin He, Juan P Rodríguez, Flavio Fontana, Matthias Faessler, Christian Forster, Jürgen Schmidhuber, Gianni Di Caro, et al. A machine learning approach to visual perception of forest trails for mobile robots. *IEEE Robotics and Automation Letters*, 1:661–667, 2015.

[98] Vern L Glaser. *Enchanted algorithms The Quantification of Organizational Decision-Making*. University of Southern California, 2014.

[99] Kurt Gray and Daniel M Wegner. Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1):125–130, 2012.

[100] Victoria Groom and Clifford Nass. Can robots be teammates?: Benchmarks in human–robot teams. *Interaction studies*, 8(3):483–500, 2007.

[101] Jaron Harambam, Dimitrios Bountouridis, Mykola Makhortykh, and Joris Van Hoboken. Designing for the better by taking users into account: A qualitative evaluation of user control mechanisms in (news) recommender systems. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 69–77, 2019.

[102] Jeffrey Heer. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116:1844–1850, 2019.

[103] Bram Hendriks, Bernt Meerbeek, Stella Boess, Steffen Pauws, and Marieke Sonneveld. Robot vacuum cleaner personality and behavior. *International Journal of Social Robotics*, 3:187–195, 2011.

[104] Martin Hilbert, Saifuddin Ahmed, Jaeho Cho, Billy Liu, and Jonathan Luu. Communicating with Algorithms: A Transfer Entropy Analysis of Emotions-based Escapes from Online Echo Chambers. *Communication Methods and Measures*, 12(4):260–275, 2018.

[105] Gunter J Hitsch, Ali Hortaçsu, and Dan Ariely. Matching and sorting in online dating. *American Economic Review*, 100(1):130–63, 2010.

[106] Shanee Honig, Alon Bartal, Yisrael Parmet, and Tal Oron-Gilad. Using online customer reviews to classify, predict, and learn about domestic robot failures. *arXiv preprint arXiv:2201.03287*, 2022.

[107] Shanee Honig and Tal Oron-Gilad. Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in psychology*, 9:861, 2018.

[108] Eric Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 159–166, 1999.

[109] Xiyang Hu, Yan Huang, Beibei Li, and Tian Lu. Uncovering the Source of Evaluation Bias in Micro-Lending. In *ICIS 2021 Proceedings*, volume 1. Association for Computing Machinery, 2021.

[110] Yin-Fu Huang and Yi-Hao Li. Sentiment translation model for expressing positive sentimental statements. In *2019 International Conference on Machine Learning and Data Engineering (iCMLDE)*, pages 79–84. IEEE, 2019.

[111] Lillian Hung, Mario Gregorio, Jim Mann, Christine Wallsworth, Neil Horne, Annette Berndt, Cindy Liu, Evan Woldum, Andy Au-Yeung, and Habib Chaudhury. Exploring the perceptions of people with dementia about the social robot paro in a hospital setting. *Dementia*, 20:485–504, 2021.

[112] Nicholas R Jennings, Luc Moreau, David Nicholson, Sarvapali Ramchurn, Stephen Roberts, Tom Rodden, and Alex Rogers. Human-agent collectives. *Communications of the ACM*, 57:80–88, 2014.

[113] Sooyeon Jeong, Cynthia Breazeal, Deirdre Logan, and Peter Weinstock. Huggable: the impact of embodiment on promoting socio-emotional interactions for young pediatric inpatients. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

[114] Neil Johnson, Guannan Zhao, Eric Hunsader, Hong Qi, Nicholas Johnson, Jing Meng, and Brian Tivnan. Abrupt rise of new machine ecology beyond human response time. *Scientific reports*, 3(1):1–7, 2013.

[115] Ece Kamar, Severin Hacker, and Eric Horvitz. Combining human and machine intelligence in large-scale crowdsourcing. In *AAMAS*, volume 12, pages 467–474, 2012.

[116] William B Kannel and Daniel L McGee. Diabetes and cardiovascular disease: the framingham study. *Jama*, 241(19):2035–2038, 1979.

[117] Serge Kernbach, Ronald Thenius, Olga Kernbach, and Thomas Schmickl. Re-embodiment of honeybee aggregation behavior in an artificial micro-robotic system. *Adaptive Behavior*, 17:237–259, 2009.

[118] Peter H Kahn, Nathan G Freier, Takayuki Kanda, Hiroshi Ishiguro, Jolina H Ruckert, Rachel L Severson, and Shaun K Kane. Design patterns for sociality in human-robot interaction. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 97–104, 2008.

[119] Hiroaki Kitano, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, and Eiichi Osawa. Robocup: The robot world cup initiative. In *Proceedings of the first international conference on Autonomous agents*, pages 340–347, 1997.

[120] Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*, 133(1):237–293, 08 2017.

[121] Jacqueline M Kory Westlund, Sooyeon Jeong, Hae W Park, Samuel Ronfard, Aradhana Adhikari, Paul L Harris, David DeSteno, and Cynthia L Breazeal. Flat vs. expressive storytelling: Young children's learning and retention of a social robot's narrative. *Frontiers in human neuroscience*, 11:295, 2017.

[122] Johannes Kunkel, Claudia Schwenger, and Jürgen Ziegler. NewsViz: Depicting and Controlling Preference Profiles Using Interactive Treemaps in News Recommender Systems. *UMAP 2020 - Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, pages 126–135, 2020.

[123] Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. Multi-agent cooperation and the emergence of (natural) language. In *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, pages 1–11, 2017.

[124] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[125] Michael L. Littman, Ifeoma Ajunwa, Guy Berger, Craig Boutilier, Morgan Currie, Finale Doshi-Velez, Gillian Hadfiel, Michael C. Horowitz, Charles Isbell, Hiroaki Kitano, Karen Levy, Terah Lyons, Melanie Mitchell, Julie Shah, Steven Sloman, Shannon Vallor, and Toby Walsh. Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report. *Stanford University, Stanford, CA*, pages 1–82, 2021.

[126] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. Talk to me: Exploring user interactions with the amazon alexa. *Journal of Librarianship and Information Science*, 51(4):984–997, 2019.

[127] Gustavo López, Luis Quesada, and Luis A Guerrero. Alexa vs. siri vs. cortana vs. google assistant: a comparison of speech-based natural user interfaces. In *International conference on applied human factors and ergonomics*, pages 241–250. Springer, 2017.

[128] Tamara Lorenz, Astrid Weiss, and Sandra Hirche. Synchrony and reciprocity: Key mechanisms for social companion robots in therapy and care. *International Journal of Social Robotics*, 8(1):125–143, 2016.

[129] John Markoff. *Machines of loving grace: The quest for common ground between humans and robots*. HarperCollins Publishers, 2016.

[130] Sean McGregor. Preventing Repeated Real World AI Failures by Cataloging Incidents: The AI Incident Database. *35th AAAI Conference on Artificial Intelligence, AAAI 2021*, 17B:15458–15463, 2021.

[131] Michelle N Meyer. Two cheers for corporate experimentation: The a/b illusion and the virtues of data-driven innovation. *Colo. Tech. LJ*, 13:273, 2015.

[132] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 220–229, 2019.

[133] Alex Mitrevski, Santosh Thoduka, Argentina Ortega Sáinz, Maximilian Schöbel, Patrick Nagel, Paul G Plöger, and Erwin Prassler. Deploying robots in everyday environments: Towards dependable and practical robotic systems. *arXiv preprint arXiv:2206.12719*, 2022.

[134] Brent Mittelstadt, Chris Russell, and Sandra Wachter. Explaining explanations in ai. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 279–288, 2019.

[135] Alexandra S Mueller, Ian J Reagan, and Jessica B Cicchino. Addressing driver disengagement and proper system use: Human factors recommendations for level 2 driving automation design. *Journal of Cognitive Engineering and Decision Making*, 15(1):3–27, 2021.

[136] Simone Natale. To believe in Siri: A critical analysis of AI voice assistants. Technical Report March, University of Bremen, 2020.

[137] Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*, pages 677–686, 2014.

[138] Yuriy Nevmyvaka, Yi Feng, and Michael Kearns. Reinforcement learning for optimized trade execution. In *Proceedings of the 23rd international conference on Machine learning*, pages 673–680, 2006.

[139] Amit Kumar Pandey and Rodolphe Gelin. A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3):40–48, 2018.

[140] Hae Won Park, Rinat Rosenberg-Kima, Maor Rosenberg, Goren Gordon, and Cynthia Breazeal. Growing growth mindset with a social robot peer. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*, pages 137–145, 2017.

[141] Rik Peeters. The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making. *Information Polity*, 25(4):507–522, 2020.

[142] Ola Pettersson. Execution monitoring in robotics: A survey. *Robotics and Autonomous Systems*, 53(2):73–88, 2005.

[143] Antonio Pérez, M Isabel García, Manuel Nieto, José L Pedraza, Santiago Rodríguez, and Juan Zamorano. Argos: An advanced in-vehicle data recorder on a massively sensorized vehicle for car driver behavior experimentation. *IEEE Transactions on Intelligent Transportation Systems*, 11(2):463–473, 2010.

[144] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. *NAACL-HLT 2016 - 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Demonstrations Session*, pages 97–101, 2016.

[145] Angel Rivas-Casado, Rafael Martinez-Tomás, and Antonio Fernández-Caballero. Multi-agent system for knowledge-based event recognition and composition. *Expert Systems*, 28(5):488–501, 2011.

[146] Florian Rosenberg and Schahram Dustdar. Design and implementation of a service-oriented business rules broker. In *Seventh IEEE International Conference on E-Commerce Technology Workshops*, pages 55–63. IEEE, 2005.

[147] Michael Rubenstein, Alejandro Cornejo, and Radhika Nagpal. Programmable self-assembly in a thousand-robot swarm. *Science*, 345(6198):795–799, 2014.

[148] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

[149] Elliot Salisbury, Ece Kamar, and Meredith Ringel Morris. Toward scalable social alt text: Conversational crowdsourcing as a tool for refining vision-to-language technology for the blind. In *Fifth AAAI Conference on Human Computation and Crowdsourcing*, 2017.

[150] Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *science*, 311(5762):854–856, 2006.

[151] Filippo Santoni de Sio and Jeroen Van den Hoven. Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, page 15, 2018.

[152] Ben Shneiderman. The limits of speech recognition. *Communications of the ACM*, 43(9):63–65, 2000.

[153] Mehdi Shanbedi, Saeed Zeinali Heris, Ahmad Amiri, Sadegh Adyani, Mohsen Alizadeh, and Majid Baniadam. Optimization of the thermal efficiency of a two-phase closed thermosyphon using active learning on the human algorithm interaction. *Numerical Heat Transfer; Part A: Applications*, 66(8):947–962, 2014.

[154] Ben Sheehan, Hyun Seung Jin, and Udo Gottlieb. Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, 115:14–24, 2020.

[155] Donghee Shin. How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance. *Computers in Human Behavior*, 109(May 2019):106344, 2020.

[156] Hirokazu Shirado and Nicholas A Christakis. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature*, 545(7654):370–374, 2017.
In this human-machine hybrid study, the authors show that simple algorithms injected into the human player can improve the results of human-to-human coordination.

[157] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[158] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[159] Wilko Schwarting, Javier Alonso-Mora, Liam Pauli, Sertac Karaman, and Daniela Rus. Parallel autonomy in automated vehicles: Safe motion generation with minimal intervention. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1928–1935. IEEE, 2017.

[160] Michel Taïx, David Flavigné, and Etienne Ferré. Human interaction with motion planning algorithm. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 67(3-4):285–306, 2012.

[161] Andreas Theodorou, Robert H Wortham, and Joanna J Bryson. Designing and implementing transparency for real time inspection of autonomous robots. *Connection Science*, 29(3):230–241, 2017.

[162] Andrea L Thomaz and Cynthia Breazeal. Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, 172(6-7):716–737, 2008.

[163] Andreas Wagner. *Robustness and evolvability in living systems*. Princeton university press, 2013.

[164] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:1606.05718*, 2016.

[165] David Watson. The rhetoric and reality of anthropomorphism in artificial intelligence. *Minds and Machines*, 29(3):417–440, 2019.

[166] John Wenskovitch, Michelle Zhou, Christopher Collins, Remco Chang, Michelle Dowling, Alex Endert, and Kai Xu. Putting the "i" in interaction: Interactive interfaces personalized to individuals. *IEEE Computer Graphics and Applications*, 40(3):73–82, 2020.

[167] Jacqueline M Kory Westlund, Hae Won Park, Randi Williams, and Cynthia Breazeal. Measuring young children's long-term relationships with social robots. In *Proceedings of the 17th ACM conference on interaction design and children*, pages 207–218, 2018.

[168] Sangseok You and Lionel Robert. Emotional attachment, performance, and viability in teams collaborating with embodied physical action (epa) robots. *You, S. and Robert, LP (2018). Emotional Attachment, Performance, and Viability in Teams Collaborating with Embodied Physical Action (EPA) Robots, Journal of the Association for Information Systems*, 19(5):377–407, 2017.

[169] Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. Learning fair representations. In *International conference on machine learning*, pages 325–333. PMLR, 2013.

[170] Xiaohua Zeng, Abraham O Fapojuwo, and Robert J Davies. Design and performance evaluation of voice activated wireless home devices. *IEEE Transactions on Consumer Electronics*, 52(3):983–989, 2006.

[171] Jie M Zhang, Mark Harman, Lei Ma, and Yang Liu. Machine learning testing: Survey, landscapes and horizons. *IEEE Transactions on Software Engineering*, 2020.

[172] Stefano Angieri, Alberto García-Martínez, Bingyang Liu, Zhiwei Yan, Chuang Wang, and Marcelo Bagnulo. An experiment in distributed internet address management using blockchains. *arXiv preprint arXiv:1807.10528*, 2018.

[173] Juliao Braga, Joao Nuno Silva, Patricia Takako Endo, Jessica Ribas, and Nizam Omar. Blockchain to improve security, knowledge and collaboration inter-agent communication over restrict domains of the internet infrastructure, with human interaction. *Brazilian Journal of Development*, 5(7):9013–9029, july 2019. DOI:10.34117/bjdv5n7-103, ISSN 2525-8761.

[174] Igor M Coelho and Vitor N Coelho. Neocompiler eco: experimentaçao de consenso em blockchain e contratos inteligentes. In *Anais do VI Workshop do testbed FIBRE*, pages 57–67. SBC, 2021.

[175] Lara Bonemer Rocha Floriani. *Smart contracts nos contratos empresariais: um estudo sobre possibilidade e viabilidade econômica de sua utilização*. Editora Dialética, 2021.

[176] Alan E. Kazdin. *The token economy: A Review and Evaluation*. Plenum Press, 2012.

[177] Antony Lewis. *The Basics of Bitcoins and Blockchains: An Introduction to Cryptocurrencies and the Technology that Powers Them*. Group, Mango Publishing, Coral Gables, FL, 1 edition, 2018.

[178] Alex Murray, Dennie Kim, and Jordan Combs. The promise of a decentralized internet: What is web 3.0 and how can firms prepare? *Business Horizons*, 65:565–570, 2022.

[179] Ankita Saxena. Workforce Diversity: A Key to Improve Productivity. *Procedia Economics and Finance*, 11:76–85, 2014.

[180] Steven D. Travers. Distributed Autonomous Organization: A Blockchain Organizational Archetype. *Strategic Management Society 37th Annual Conference*, 2017.

[181] Shermin Voshmgir. *Economia dos Tokens: Como a Web3 está reinventando a internet e a relação entre os agentes econômicos*. Token Kitchen, 2 edition, 2020.

[182] Aries Wanlin Wang. *Crypto Economy: How Blockchain, Cryptocurrency and Token-Economy are Disrupting the Financial World*. Skyhorse Publishing, 2018.

[183] Guy R Vishnia and Gareth W Peters. Auditchain: A trading audit platform over blockchain. *Frontiers in Blockchain*, 3:9, 2020.

[184] W Brian Arthur. Complexity and the economy. In *Handbook of Research on Complexity*. Edward Elgar Publishing, 2009.

[185] Nicolas Bouleau. On excessive mathematization, symptoms, diagnosis and philosophical bases for real world knowledge. *Real World Economics*, 57:90–105, 2011.

[186] Kyle Croman, Christian Decker B, Ittay Eyal, Adem Efe Gencer, Ari Juels, Ahmed Kosba, Andrew Miller, Prateek Saxena, Dawn Song, and Roger Wattenhofer. On Scaling Decentralized Blockchains (A Position Paper). *Lecture Notes in Computer Science*, 9604:106–125, 2016.

[187] Eric Budish, Peter Cramton, and John Shim. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4):1547–1621, 2015.

[188] John Cartlidge, Marco De Luca, Charlotte Szostek, and Dave Cliff. Too fast too furious: faster financial-market trading agents can give less efficient markets. In *ICAART-2012: 4th International Conference on Agents and Artificial Intelligence*, pages 126–135. SciTePress, 2012.

[189] Le Chen and Christo Wilson. Observing algorithmic marketplaces in-the-wild. *ACM SIGecom Exchanges*, 15(2):34–39, 2017.

[190] Le Chen, Alan Mislove, and Christo Wilson. An empirical analysis of algorithmic pricing on amazon marketplace. In *Proceedings of the 25th international conference on World Wide Web*, pages 1339–1349, 2016.

[191] J Doyne Farmer and Spyros Skouras. An ecological perspective on the future of computer trading. *Quantitative Finance*, 13:325–346, 2013.

[192] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59:96–104, 2016.

[193] Michael Kearns, Alex Kulesza, and Yuriy Nevmyvaka. Empirical limitations on high-frequency trading profitability. *The Journal of Trading*, 5:50–62, 2010.

[194] Andrei A Kirilenko and Andrew W Lo. Moore's law versus murphy's law: Algorithmic trading and its discontents. *Journal of Economic Perspectives*, 27(2):51–72, 2013.

[195] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent trade-offs in the fair determination of risk scores. *Leibniz International Proceedings in Informatics, LIPIcs*, 67:1–23, 2017.

[196] Jon Kleinberg and Sigal Oren. Mechanisms for (mis) allocating scientific credit. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 529–538, 2011.

[197] Farshad Kooti, Mihajlo Grbovic, Luca Maria Aiello, Nemanja Djuric, Vladan Radosavljevic, and Kristina Lerman. Analyzing uber's ride-sharing economy. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 574–582, 2017.

[198] Jaron Lanier. *You are not a gadget*. Vintage, 2010.

[199] Michael Latzer, Katharina Hollnbuchner, Natascha Just, and Florian Saurwein. The economics of algorithmic selection on the internet. In *Handbook on the Economics of the Internet*, pages 395–425. University of Zurich, 2014.

[200] Albert J Menkveld. The economics of high-frequency trading. *Annual Review of Financial Economics*, 8:1–24, 2016.

[201] Miriam Naigembe. *Bank lending policy, credit scoring and the survival of Loans: A case study of banks X and Y*. PhD thesis, Makerere University, 2010.

[202] David C Parkes and Michael P Wellman. Economic reasoning and artificial intelligence. *Science*, 349(6245):267–272, 2015.

[203] Frank Pasquale. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Harvard University Press, 2015.

[204] Kasper Roszbach. Bank lending policy, credit scoring, and the survival of loans. *Review of Economics and Statistics*, 86(4):946–958, 2004.

[205] Jonathan JJM Seddon and Wendy L Currie. A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research*, 70:300–307, 2017.

[206] Robert J Shiller. *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press, 2020.

[207] Chih-Fong Tsai and Jhen-Wei Wu. Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert systems with applications*, 34(4):2639–2649, 2008.

[208] Michael P. Wellman, Peter R. Wurman, Kevin O'Malley, Roshan Bangera, Daniel Reeves, William E Walsh, et al. Designing the market game for a trading agent competition. *IEEE Internet Computing*, 5(2):43–51, 2001.

[209] Alexandre Aronne, Aureliano Bressan, and Haroldo Guimarães Brasil. *Mensuração e Gerenciamento de Riscos Corporativos: Aplicações de Cash Flow at Risk e Real Options*. Saint Paul Editora, 2021.

[210] Rogério Silva Nacif. *Operações Eficientes, Empresas Rentáveis: Melhorando os Resultados Financeiros por Meio da Gestão de Operações*. Aquila, 2021.

[211] Carlo Appugliese, Paco Nathan, and William S Roberts. *Agile AI: A Practical Guide to Building AI Applications and Teams*. O'Reilly, 2020.

[212] Kenneth Appel, Wolfgang Haken, and John Koch. Every planar map is four colorable. part ii: Reducibility. *Illinois Journal of Mathematics*, 21(3):491–567, 1977.

[213] Kenneth Appel and Wolfgang Haken. Every planar map is four colorable. *Bulletin of the American mathematical Society*, 82(5):711–712, 1976.

[214] Per Bak, Kan Chen, and Michael Creutz. Self-organized criticality in the game of life. *Nature*, 342(6251):780–782, 1989.

[215] Marc G Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*, 47:253–279, 2013.

[216] Roger Bemelmans, Gert Jan Gelderblom, Pieter Jonker, and Luc De Witte. Socially assistive robots in elderly care: a systematic review into effects and effectiveness. *Journal of the American Medical Directors Association*, 13(2):114–120, 2012.

[217] Andrew Berdahl, Colin J Torney, Christos C Ioannou, Jolyon J Faria, and Iain D Couzin. Emergent sensing of complex environments by mobile animal groups. *Science*, 339(6119):574–576, 2013.

[218] Ana Berdasco, Gustavo López, Ignacio Diaz, Luis Quesada, and Luis A Guerrero. User experience comparison of intelligent personal assistants: Alexa, google assistant, siri and cortana. *Multidisciplinary Digital Publishing Institute Proceedings*, 31(1):51, 2019.

[219] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First monday*, 21(11-7), 2016.

[220] Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Communications of the ACM*, 60(11):81–88, 2017.

[221] Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. Deep blue. *Artificial intelligence*, 134(1-2):57–83, 2002.

[222] Juan Miguel Carrascosa, Jakub Mikians, Ruben Cuevas, Vijay Erramilli, and Nikolaos Laoutaris. I always feel like somebody's watching me: measuring online behavioural advertising. In *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies*, pages 1–13, 2015.

[223] Maurice Chiodo and Toby Clifton. The importance of ethics in mathematics. *European Mathematical Society Magazine*, 114:34–37, 2019.

[224] BACEN. LIFT Challenge, 2022. Acessed in 03/09/2022.

[225] Wasifa Chowdhury. *Employing neural hierarchical model with pointer generator networks for abstractive text summarization*. PhD thesis, Simon Frazer University: School of Computing Science, 2019.

[226] Battista Biggio, Igino Corona, Davide Maiorca, Blaine Nelson, Nedim Šrndić, Pavel Laskov, Giorgio Giacinto, and Fabio Roli. Evasion attacks against machine learning at test time. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 387–402. Springer, 2013.

[227] Jack Clark and Ray Perrault. Introduction to the AI index report 2022. Technical report, Stanford University, 2022.

[228] Amit Datta, Michael Carl Tschantz, and Anupam Datta. Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *arXiv preprint arXiv:1408.6491*, 2014.

[229] Yue Deng, Feng Bao, Youyong Kong, Zhiquan Ren, and Qionghai Dai. Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3):653–664, 2016.

[230] Pedro Domingos. *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books, 2015.

[231] Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017.

[232] Sebastian Elbaum and John C Munson. Software black box: an alternative mechanism for failure analysis. In *Proceedings 11th International Symposium on Software Reliability Engineering. ISSRE 2000*, pages 365–376. IEEE, 2000.

[233] Giuliano Da Empoli. *Os engenheiros do caos*. Vestígio, Belo Horizonte, 1 edition, 2019.

[234] Amir Globerson and Sam Roweis. Nightmare at test time: robust learning by feature deletion. In *Proceedings of the 23rd international conference on Machine learning*, pages 353–360, 2006.

[235] Dennis R Grossi. Aviation Recorder Overview. In *International Symposium On Transportation Recorders*, page 12, 2006.

[236] Jose Hernandez-Orallo. Beyond the turing test. *Journal of Logic, Language and Information*, 9(4):447–466, 2000.

[237] John Kay and Mervyn King. *Radical uncertainty: Decision-making beyond the numbers*. WW Norton & Company, 2020.

[238] Himabindu Lakkaraju, Ece Kamar, Rich Caruana, and Eric Horvitz. Identifying unknown unknowns in the open world: Representations and policies for guided exploration. In *Thirty-first aaai conference on artificial intelligence*, 2017.

[239] Tian-Shyug Lee and I-Fei Chen. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with applications*, 28(4):743–752, 2005.

[240] Joel Z Leibo, Cyprien de Masson d'Autume, Daniel Zoran, David Amos, Charles Beattie, Keith Anderson, Antonio García Castañeda, Manuel Sanchez, Simon Green, Audrunas Gruslys, et al. Psychlab: a psychology laboratory for deep reinforcement learning agents. *arXiv preprint arXiv:1801.08116*, 2018.

[241] Nancy G. Leveson. *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, 2011.

[242] C Dianne Martin. The myth of the awesome thinking machine. *Communications of the ACM*, 36(4):120–133, 1993.

[243] Simone Natale et al. *Deceitful media: Artificial intelligence and social life after the Turing test*. Oxford University Press, USA, 2021.

[244] Olfa Nasraoui and Patrick Shafto. Human-Algorithm Interaction Biases in the Big Data Cycle: A Markov Chain Iterated Learning Framework. *arXiv*, 2016.

[245] Randolph M Nesse. Tinbergen's four questions, organized: a response to bateson and laland. *Trends in Ecology & Evolution*, 28(12):681–82, 2013.

[246] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318, 2002.

[247] Judea Pearl and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, first edition, 2018.

[248] David L. Poole and Alan K. Mackworth. *Artificial Intelligence: foundations of computational agents*. Cambridge University Press, second edition, 2017.

[249] David L Poole and Alan K Mackworth. *Artificial Intelligence: foundations of computational agents*. Cambridge University Press, 2010.

[250] Stuart Russel and Peter Norvig. *Artificial Intelligence*. Prentice Hall, New York, 3 edition, 2010.

[251] Jonathan Schaeffer, Neil Burch, Yngvi Bjornsson, Akihiro Kishimoto, Martin Muller, Robert Lake, Paul Lu, and Steve Sutphen. Checkers is solved. *science*, 317(5844):1518–1522, 2007.

[252] Greg Siegel. *Forensic media: Reconstructing accidents in accelerated modernity*. Duke University Press, 2014.

[253] Robert Skidelsky. *Information retrieval and hypertext*. Yale University Press, 2020.

[254] Daniel Smilkov, Nikhil Thorat, Been Kim, Fernanda Viégas, and Martin Wattenberg. Smoothgrad: removing noise by adding noise. *arXiv preprint arXiv:1706.03825*, 2017.

[255] Megha Srivastava, Hoda Heidari, and Andreas Krause. Mathematical notions vs. human perception of fairness: A descriptive approach to fairness for machine learning. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2459–2468, 2019.

[256] Venkatramanan S Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. The darpa twitter bot challenge. *Computer*, 49(6):38–46, 2016.

[257] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*, pages 1–10, 2014.

[258] Florian Tramèr, Alexey Kurakin, Nicolas Papernot, Ian Goodfellow, Dan Boneh, and Patrick McDaniel. Ensemble adversarial training: Attacks and defenses. In *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pages 1–22, 2018.

[259] Milena Tsvetkova, Taha Yasseri, Eric T. Meyer, J. Brian Pickering, Vegard Engen, Paul Walland, Marika Lüders, Asbjørn Følstad, and George Bravos. Understanding Human-Machine Networks. *ACM Computing Surveys*, 50(1):1–35, 2018.

[260] Milena Tsvetkova, Ruth García-Gavilanes, Luciano Floridi, and Taha Yasseri. Even good bots fight: The case of wikipedia. *PloS one*, 12(2):e0171774, 2017.

[261] Alan M Turing. Computing machinery and intelligence. *Mind*, LIX(236):433–460, 1950.

[262] Koen Van De Sande, Theo Gevers, and Cees Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2009.

[263] Carissa Véliz. *Privacy is power*. Melville House, 2021.

[264] Xingyu Xing, Wei Meng, Dan Doozan, Alex C Snoeren, Nick Feamster, and Wenke Lee. Take this personally: Pollution attacks on personalized services. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 671–686, 2013.

[265] Yu Yao and Ella Atkins. The smart black box: A value-driven high-bandwidth automotive event data recorder. *IEEE Transactions on Intelligent Transportation Systems*, 22(3):1484–1496, 2020.

[266] Quan-shi Zhang and Song-Chun Zhu. Visual interpretability for deep learning: a survey. *Frontiers of Information Technology & Electronic Engineering*, 19(1):27–39, 2018.

[267] Zhiming Zhou, Han Cai, Shu Rong, Yuxuan Song, Kan Ren, Weinan Zhang, Yong Yu, and Jun Wang. Activation maximization generative adversarial nets. *arXiv preprint arXiv:1703.02000*, 2017.

[268] Susmit Jha, Tuhin Sahai, Vasumathi Raman, Alessandro Pinto, and Michael Francis. Explaining ai decisions using efficient methods for learning sparse boolean formulae. *Journal of Automated Reasoning*, 63(4):1055–1075, 2019.

[269] Patrícia Gomes Rêgo de Almeida. Regulação da Inteligência Artificial: Ação Coletiva que Requer Governança. *Computação Brasil*, 7:23–26, 2022.

[270] Margaret Boden, Joanna Bryson, Darwin Caldwell, Kerstin Dautenhahn, Lilian Edwards, Sarah Kember, Paul Newman, Vivienne Parry, Geoff Pegman, Tom Rodden, Tom Sorrell, Mick Wallis, Blay Whitby, and Alan Winfield. Principles of robotics: regulating robots in the real world. *Connection Science*, 29(2):124–129, 2017.

[271] Kevin Bonsor and Nathan Chandler. How black boxes work. *HowStuffWorks, June*, 13, 2001.

[272] Fernanda Braganca and Renata Braga. Os Desafios da Regulamentação Jurídica da Inteligência Artificcial no Brasil. *Computação Brasil*, 7:19–22, 2022.

[273] Lindell Bromham, Russell Dinnage, and Xia Hua. Interdisciplinary research has consistently lower funding success. *Nature*, 534(7609):684–687, 2016.

[274] Joanna J Bryson, Mihailis E Diamantis, and Thomas D Grant. Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence and Law*, 25(3):273–291, 2017.

[275] Roger Clarke. Regulatory alternatives for ai. *Computer Law & Security Review*, 35(4):398–409, 2019.

[276] Europe Commission. Coordinated Plan on Artificial Intelligence 2021 Review. Technical report, European Commission, Brussels, 2021.

[277] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pages 797–806, 2017.

[278] Iain D Couzin, Christos C Ioannou, Güven Demirel, Thilo Gross, Colin J Torney, Andrew Hartnett, Larissa Conradt, Simon A Levin, and Naomi E Leonard. Uninformed individuals promote democratic consensus in animal groups. *science*, 334(6062):1578–1580, 2011.

[279] Kate Crawford, Meredith Whittaker, Madeleine Clare Elish, Solon Barocas, Aaron Plasek, and Kadija Ferryman. The ai now report. *The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term*, 2016.
O paper analisou a IA em relação a quatro temas-chave: Saúde, Trabalho, Desigualdade e Ética. E oferece oito (8) recomendações como passos práticos que os interessados envolvidos em vários pontos na produção, uso, governança e avaliação dos sistemas de IA poderiam tomar para enfrentar os desafios e oportunidades de curto prazo criados pela rápida implementação de IA através dos domínios social e econômico.

[280] Kenneth Cukier, Viktor Mayer-Schönberger, and Francis de Véricourt. *Framers: Human advantage in an age of technology and turmoil*. Penguin, 2022.

[281] Harrison Edwards and Amos Storkey. Censoring representations with an adversary. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, pages 1–14, 2016.

[282] EPI. Algorithmic transparency: End secret profiling. Technical report, Electronic Privacy Information Center, 2015.

[283] Virginia Eubanks. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press, 2018.

[284] Gregory Falco, Ben Shneiderman, Julia Badger, Ryan Carrier, Anton Dahbura, David Danks, Martin Eling, Alwyn Goodloe, Jerry Gupta, Christopher Hart, et al. Governing ai safety through independent audits. *Nature Machine Intelligence*, 3(7):566–571, 2021.

[285] Hanming Fang and Andrea Moro. Theories of statistical discrimination and affirmative action: A survey. In *Handbook of Social Economics*, volume 1, pages 133–200. Elsevier B.V., 2011.

[286] Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for ai. *Berkman Klein Center Research Publication*, 2020-1, 2020.

[287] Ana Frazão. Discriminação Algoritmica: a relação entre homens e máquinas. Coluna Jota, junho 2021. Trabalho dividido em treze partes.

[288] Future of Live Institute. Autonomous Weapons: An Open Letter from AI and Robotics Reserarchers. `https://futureoflife.org/2016/02/09/open-letter-autonomous-weapons-ai-robotics/?cn-reloaded=1&cn-reloaded=1`, July 2018.

[289] IEEE. Ethically Aligned Design: Version 2 - For Public Discussion. *IEEE Standards*, pages 1–263, 2017.

[290] CEI. Our vision. Technical report, Council on ExtendedIntelligence, 2022.

[291] Lucas D. Introna and Helen Nissenbaum. Shaping the web: Why the politics of search engines matters. *Computer Ethics*, pages 157–173, 2017.

[292] Joosr. *A Joosr Guide to... Weapons of Math Destruction by Cathy O'Neil: How Big Data Increases Inequality and Threatens Democracy*. Broadway books, 2016.

[293] Daniel Kahneman, Olivier Sibony, and CR Sunstein. *Noise*. HarperCollins UK, 2022.

[294] Daniel Kahneman, AM Rosenfield, L Gandhi, and T Blaser. Noise: How to overcome the high. *Havard Business Review*, 2016. Available in https://hbr.org/2016/10/noise.

[295] Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[296] Pratyusha Kalluri et al. Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*, 583:169–169, 2020.

[297] Krishna M. Kavi. Beyond the Black Box. *IEE Spectrum*, 47(8):46—-51, 2021.

[298] Pranav Khadpe, Ranjay Krishna, Li Fei-Fei, Jeffrey T Hancock, and Michael S Bernstein. Conceptual metaphors impact perceptions of human-ai collaboration. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–26, 2020.

[299] Mervyn King and John Kay. *Radical Uncertainty: Decision-making for an unknowable future*. Hachette UK, 2020.

[300] Nicole C Krämer, Astrid von der Pütten, and Sabrina Eimler. Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In *Human-computer interaction: The agency perspective*, pages 215–240. Springer, 2012.

[301] Samantha Krening and Karen M Feigh. Interaction Algorithm Effect on Human Experience. *ACMTrans. Human-Robot Interact*, 7(2):22, 2018.

[302] Armin Krishnan. *Killer robots: legality and ethicality of autonomous weapons*. Routledge, 2016.

[303] David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani. The parable of google flu: traps in big data analysis. *Science*, 343(6176):1203–1205, 2014.

[304] Heidi Ledford. Team science. *Nature*, 525(7569):308–311, 2015.

[305] Joseph CR Licklider. Man-computer symbiosis. *IRE transactions on human factors in electronics*, pages 4–11, 1960.

[306] Claudia Bauzer Medeiros. Dados, Algoritmos, Máquinas e Pessoas. *Computação Brasil*, 7:11–14, 2022.

[307] David A Mindell. *Our robots, ourselves: Robotics and the myths of autonomy*. Viking, 2015.

[308] Sendhil Mullainathan. Biased algorithms are easier to fix than biased people. *The New York Times*, 2019.

[309] NSCAI. National Security Commission on Artificial Intelligence - Interim Report. *National Security Commission on Artificial Intelligence Report*, pages 1–101, 2019.

[310] Cathy O'neil. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway books, 2016.

[311] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[312] Cristina Godoy Bernardo de Oliveira, João Paulo Cândia Veiga, and Fabio G. Cozman. Regulação da Inteligência Artificial: Qual o Modelo Adotar. *Computação Brasil*, 7:28–31, 2022.

[313] Cathy O'Neil, editor. *Algoritmos de Destruição em Massa*. Editora Rua do Sabão, Santo Andrá, SP, 2020.

[314] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 30(3):286–297, 2000.

[315] Frank Pasquale. *New laws of robotics: defending human expertise in the age of AI*. Belknap Press, 2020.

[316] Komal Patel. Testing the Limits of the First Amendment: How a CFAA Prohibition on Online Antidiscrimination Testing Infringes on Protected Speech Activity. *SSRN Electronic Journal*, pages 1–46, 2017.

[317] Walt L Perry. *Predictive policing: The role of crime forecasting in law enforcement operations*. Rand Corporation, 2013.

[318] Sundar Pichai. AI at Google: our principles. *2018-07-07*, page 1, 2018.

[319] Magaly Prado. *Fake news e inteligência artificial: Oo poder dos algoritmos na guerra da desinformação*, volume 1. Almedina Brasil, 2022.

[320] Tony J Prescott and Julie M Robillard. Are friends electric? the benefits and risks of human-robot relationships. *Iscience*, 24(1):101993, 2021.

[321] Iyad Rahwan, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W Crandall, Nicholas A Christakis, Iain D Couzin, Matthew O Jackson, et al. Machine behaviour. *Nature*, 568(7753):477–486, 2019.
Trata do comportamento dos sistemas de AI, cuja compreensão é essencial para nossa capacidade de controlar suas ações, colher seus benefícios e minimizar seus danos. Em síntese, entender o comportamento dos sistemas de AI, para maximizar seus benefícios e minimizar os danos em relação à humanidade.

[322] Byron Reeves and Clifford Nass. How people treat computers, television, and new media like real people and places, 1996.

[323] Tahira Reid and James Gibert. Inclusion in human–machine interactions. *Science*, 375(6577):149–150, 2022.

[324] Lionel P Robert. The Growing Problem of Humanizing Robots. *International Robotics & Automation Journal*, 3(1):1–2, 2017.

[325] Margaret E Roberts. Censored. In *Censored*. Princeton University Press, 2018.

[326] W Teed Rockwell. Algorithms and stories. *Human Affairs*, 23(4):633–644, 2013.

[327] Heather M Roff. The strategic robot problem: Lethal autonomous weapons in war. *Journal of Military Ethics*, 13(3):211–227, 2014.

[328] Stuart Russell. *Human compatible: Artificial intelligence and the problem of control*. Penguin, 2019.

[329] Jário Santos, Ig Bittencourt, Marcelo Reis, Geiser Chalco, and Seiji Isotani. Two billion registered students affected by stereotyped educational environments: an analysis of gender-based color bias. *HUMANITIES AND SOCIAL SCIENCES COMMUNICATIONS* |, 9(249):1—-16, 2022.

[330] Daniel Schiff, Aladdin Ayesh, Laura Musikanski, and John C Havens. Ieee 7010: A new standard for assessing the well-being implications of artificial intelligence. In *2020 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 2746–2753. IEEE, 2020.

[331] Helen Sharp, Yvonne Rogers, and Jennifer Preece. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons Inc, fifth edition, 2019.

[332] Ben Shneiderman. Design lessons from ai's two grand goals: Human emulation and useful applications. *IEEE Transactions on Technology and Society*, 1(2):73–82, 2020.

[333] Ben Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6):495–504, 2020.

[334] Ben Shneiderman. The dangers of faulty, biased, or malicious algorithms requires independent oversight. *Proceedings of the National Academy of Sciences*, 113(48):13538–13540, 2016.

[335] Ben Shneiderman, Catherine Plaisant, Maxine S Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

[336] Ben Shneiderman. Human responsibility for autonomous agents. *IEEE intelligent systems*, 22(2):60–61, 2007.

[337] Ben Shneiderman and Pattie Maes. Direct manipulation vs. interface agents. *interactions*, 4(6):42–61, 1997.

[338] Ben Shneiderman. The future of interactive systems and the emergence of direct manipulation. *Behaviour & Information Technology*, 1(3):237–256, 1982.

[339] Ben Shneiderman. Direct manipulation: A step beyond programming languages. In *Proceedings of the Joint Conference on Easier and More Productive Use of Computer Systems.(Part-II): Human Interface and the User Interface-Volume 1981*, page 143, 1981.

[340] Marlo Souza. Tecnologias da Linguagem, Ética em IA e Regulamentação. *Computação Brasil*, 7:32–35, 2022.

[341] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee (Anno) Saxenian, Julie Shah, Milind Tambe, and Astro Teller. Artificial Intelligence and Life in 2030: the one hundred year study on artificial intelligence. Technical report, Stanford University, September 2016.

[342] Megan K Strait, Cynthia Aguillon, Virginia Contreras, and Noemi Garcia. The public's perception of humanlike robots: Online social commentary reflects an appearance-based uncanny valley, a general fear of a "technology takeover", and the unabashed sexualization of female-gendered robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 1418–1423. IEEE, 2017.

[343] Barry Strauch. Ironies of automation: Still unresolved after all these years. *IEEE Transactions on Human-Machine Systems*, 48(5):419–433, 2017.

[344] Wenlong Sun, Olfa Nasraoui, and Patrick Shafto. *Evolution and impact of bias in human and machine learning algorithm interaction*, volume 15. Public Library of Science San Francisco, CA USA, 2020.

[345] Latanya Sweeney. Discrimination in online ad delivery. *Communications of the ACM*, 56(5):44–54, 2013.

[346] Niko Tinbergen. On aims and methods of ethology. *Animal Biology*, 55(4):297–321, 2005.

[347] Ufuk Topcu, Nadya Bliss, Nancy Cooke, Missy Cummings, Ashley Llorens, Howard Shrobe, and Lenore Zuck. Assured autonomy: Path toward living with autonomous systems we can trust. *arXiv preprint arXiv:2010.14443*, 2020.

[348] Zeynep Tufekci. Youtube, the great radicalizer. *The New York Times*, 10(3):2018, 2018.

[349] Zeynep Tufekci. Engineering the public: Big data, surveillance and computational politics. *First Monday*, 2014.

[350] UN. Report of the Working Group on Internet Governance. Technical report, United Nations, 2005.

[351] University of Montreal. Montréal Declaration for a Responsible Development of Artificial Intelligence. Technical report, Universidade de Montreal, 2018.

[352] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *science*, 359(6380):1146–1151, 2018.

[353] Paul Voosen. The ai detectives, 2017.

[354] Sara Wachter-Boettcher. *Technically wrong: Sexist apps, biased algorithms, and other threats of toxic tech*. WW Norton & Company, 2017.

[355] Chathurika S Wickramasinghe, Daniel L Marino, Javier Grandio, and Milos Manic. Trustworthy ai development guidelines for human system interaction. In *2020 13th International Conference on Human System Interaction (HSI)*, pages 130–136. IEEE, 2020.

[356] Christine T. Wolf and Jeanette L. Blomberg. Evaluating the promise of human-algorithm collaborations in everyday work practices. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–23, 2019.

[357] David D Woods, James Tittle, Magnus Feil, and Axel Roesler. Envisioning human-robot coordination in future operations. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2):210–218, 2004.

[358] Wei Xu. Toward human-centered ai: a perspective from human-computer interaction. *interactions*, 26(4):42–46, 2019.

[359] Sergio Amadeu Silveira. Governo dos algoritmos. *Revista de Políticas Públicas*, 21(1):267–281, 2017.

[360] Can Yavuz. *Machine Bias Artificial Intelligence and Discrimination*. PhD thesis, Lund University, 2019.

[361] Shunyuan Zhang, Nitin Mehta, Param Vir Singh, and Kannan Srinivasan. Can an ai algorithm mitigate racial economic inequality? an analysis in the context of airbnb. *An Analysis in the Context of Airbnb (January 21, 2021). Rotman School of Management Working Paper*, 2021.

[362] Ignas Kalpokas. *Algorithmic Governance*, volume 9. Palgrave Macmillan, 2019.

[363] Meredith Broussard. *Artificial unintelligence: How computers misunderstand the world*. MIT Press, 2018.

[364] César A Hidalgo, Diana Orghian, Jordi Albo Canals, Filipa De Almeida, and Natalia Martín. *How humans judge machines*. MIT Press, 2021.

[365] Alexis Lambert, Nahal Norouzi, Gerd Bruder, and Gregory Welch. A systematic review of ten years of research on human interaction with social robots. *International Journal of Human–Computer Interaction*, 36(19):1804–1817, 2020.

[366] Yi Mou, Changqian Shi, Tianyu Shen, and Kun Xu. A systematic review of the personality of robot: Mapping its conceptualization, operationalization, contextualization and effects. *International Journal of Human–Computer Interaction*, 36(6):591–605, 2020.

[367] Sarah Sebo, Brett Stoll, Brian Scassellati, and Malte F Jung. Robots in groups and teams: a literature review. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–36, 2020.

[368] Weiyu Wang and Keng Siau. Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda. *Journal of Database Management (JDM)*, 30(1):61–79, 2019.

[369] Cigdem BAŞFIRINCI and Zuhal ÇİLİNGİR. Anthropomorphism and advertising effectiveness: Moderating roles of product involvement and the type of consumer need. *Journal of Social and Administrative Sciences*, 2(3):108–131, 2015.

[370] Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot, editors. *Media Technologies: Essays on Communication, Materiality, and Society*. The MIT Press, 2014.

[371] Mark Granovetter and Roland Soong. Threshold models of diffusion and collective behavior. *Journal of Mathematical sociology*, 9:165–179, 1983.

[372] Mark Granovetter. Threshold models of collective behavior. *American journal of sociology*, 83:1420–1443, 1978.

[373] Anikó Hannák, Claudia Wagner, David Garcia, Alan Mislove, Markus Strohmaier, and Christo Wilson. Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1914–1933, 2017.

[374] Leila Hudson, Colin S Owens, and Matt Flannes. Drone warfare: Blowback from the new american way of war. *Middle East Policy*, 18:122–132, 2011.

[375] Colm Kearns, Gary Sinclair, Jack Black, Mark Doidge, Thomas Fletcher, Daniel Kilvington, Katie Liston, Theo Lynn, and Pierangelo Rosati. A scoping review of research on online hate and sport. *Communication & Sport*, pages 1–29, 2022.

[376] Peter M Krafft, Michael Macy, and Alex" Sandy" Pentland. Bots as virtual confederates: design and ethics. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 183–190, 2017.

[377] Thomas S. Kuhn. *A Estrutura das Revoluções Científicas*. Perspectiva, São Paulo, 1 edition, 1996.

[378] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. Computational social science. *Science*, 323(5915):721–723, 2009.

[379] Yuhua Liang and Seungcheol Austin Lee. Fear of autonomous robots and artificial intelligence: Evidence from national representative data with probability sampling. *International Journal of Social Robotics*, 9(3):379–384, 2017.

[380] C Dianne Martin. Eniac: press conference that shook the world. *IEEE Technology and Society Magazine*, 14(4):3–10, 1995.

[381] Neil McBride. Robot enhanced therapy for autistic children: An ethical analysis. *IEEE Technology and Society Magazine*, 39(1):51–60, 2020.

[382] Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, and Sune Lehmann. Evidence of complex contagion of information in social media: An experiment using twitter bots. *PloS one*, 12(9):e0184148, 2017.

[383] Lewis Mumfordt. *Technics and Civilization*. University of Chicago Press, Chicago, 1934.

[384] Kathleen Richardson, Mark Coeckelbergh, Kutoma Wakunuma, Erik Billing, Tom Ziemke, Pablo Gomez, Bram Vanderborght, and Tony Belpaeme. Robot enhanced therapy for children with autism (dream): A social model of autism. *IEEE Technology and society magazine*, 37(1):30–39, 2018.

[385] Michael P Wellman and Uday Rajan. Ethical issues for autonomous trading agents. *Minds and Machines*, 27(4):609–624, 2017.

[386] Marc Wiedermann, E Keith Smith, Jobst Heitzig, and Jonathan F Donges. A network-based microfoundation of granovetter's threshold model for social tipping. *Scientific reports*, 10(1):1–10, 2020.

[387] Alan FT Winfield and Marina Jirotka. The case for an ethical black box. In *Annual Conference Towards Autonomous Robotic Systems*, pages 262–273. Springer, 2017.

[388] Alan FT Winfield and Marina Jirotka. Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133):20180085, 2018.

[389] Néstor García Canclini. *Ciudadanos reemplazados por algoritmos*. Calas, 2019.

[390] Joseph Weizenbaum. *Computer Power and Human Reason: From Judgement to Calculation*. W. H. Freeman and Company, 1976.

[391] Juliao Braga, Francisco Regateiro, Itana Stiubiener, and Juliana C Braga. Project for the development of a paper on algorithm and data governance. Portuguese version: `https://osf.io/xcpsd`, Sep 2022.

[392] Juliao Braga, Francisco Regateiro, Itana Stiubiener, and Juliana C Braga. Human-algorithm: Governance, Sep 2022. DOI; 10.17605/OSF.IO/K35FV.

[393] Ritesh Modi. *Solidity Programming Essentials: A beginner's guide to build smart contracts for Ethereum and blockchain*. Packt Publishing Ltd, 2018.

[394] Juliao Braga, Francisco Regateiro, and Itana Stiubiener. Human-algorithm (private), Sep 2022. Acessed in 03/09/2022.