

novembro/2020 • n. 43

COMPUTAÇÃO

REVISTA DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO  BRASIL



ARTIFICIALMENTE HUMANO OU HUMANAMENTE ARTIFICIAL?

Desafios da Sociedade 5.0

EDITORIAL

CIÊNCIA E TECNOLOGIAS DA COMPUTAÇÃO COM RESPONSABILIDADE SOCIOAMBIENTAL



RAIMUNDO JOSÉ DE ARAÚJO MACÊDO

Presidente da Sociedade Brasileira de Computação (SBC)

O surgimento e desenvolvimento da ciência da computação e suas tecnologias, desde a primeira metade do século XX até os dias atuais, mudou drasticamente os meios de produção de riquezas, revolucionou a comunicação, o entretenimento, a saúde, a educação e aproximou países e culturas. Assim como ocorreu com outros instrumentos de desenvolvimento, como máquinas a vapor, lâmpadas elétricas e comunicação por ondas de rádio, a computação, de uma forma ainda mais radical, pôde e pode acelerar processos sócio-econômico-ambientais, sejam eles benéficos ou maléficos. Portanto, a computação precisa ser trabalhada com o devido cuidado para que os benefícios por ela produzidos sejam amplos e para todos, sem discriminar povos, etnias, gêneros, nações, ou causar prejuízos ambientais, ou sociais. A relevância dessas tecnologias no desenvolvimento da humanidade está bem colocada em declaração recente de um dos maiores líderes religiosos do planeta, o Papa Francisco, pedindo aos fiéis “que rezem para que os avanços de tecnologias como inteligência artificial e robótica possam ajudar a diminuir a desigualdade entre os povos e servir à humanidade”.

A SBC, desde sua criação há 42 anos, tem assumido seu papel no desenvolvimento da computação com responsabilidade socioambiental. Coerente com esses princípios e com os desafios contemporâneos, trouxe o tema “Sociedade 5.0” para seu Congresso anual de 2020, em Cuiabá, no Mato Grosso. O termo “Sociedade 5.0” se refere à utilização das novas tecnologias baseadas na inteligência artificial, big data, Internet das Coisas, robótica, entre outras. A extensa programação de nosso Congresso de Cuiabá explora vários ângulos dessas

questões, desde a adequada educação para a formação de profissionais, a princípios éticos, passando pelo combate à desinformação e notícias falsas, necessária sustentabilidade ambiental, entre outros aspectos da vida humana e do planeta.

O país precisa se preparar para se tornar um ator mundial importante na produção de tecnologias da computação, para isso é necessário reforçar o sistema de educação e pesquisa científica na área. Nossas conferências, simpósios, workshops, escolas regionais, e suas publicações, são instrumentos que potencializam o desenvolvimento brasileiro nesse campo. Visando o fortalecimento de nosso sistema de difusão da informação científica e tecnológica, propusemos, debatemos e aprovamos, no Conselho da SBC, uma diretriz para que todos os nossos eventos realizem suas publicações no modelo aberto e na SOL, nossa biblioteca digital. Instituímos um código de conduta para autores de publicações da SBC, com amplo debate interno, consulta pública e aprovação em assembleia extraordinária. Criamos um periódico, o SBC Reviews, inteiramente dedicado a revisões sistemáticas de literatura, que acreditamos será de grande relevância para nossos estudantes em atividades de pesquisa. Essas iniciativas somam-se a tantas outras, tradicionais ou recentes, que colocam nossa sociedade cada vez mais no caminho que buscamos: o desenvolvimento socioeconômico do país, inclusivo e com respeito ao meio ambiente e à vida.

Este número da revista Computação Brasil traz artigos que refletem bem as discussões de nosso Congresso anual, além de dois artigos de autoria dos ganhadores do Prêmio Mérito Científico da SBC de 2020. Agradecemos aos autores e aos editores deste número especial de nossa revista.

novembro/2020 • n. 43

COMPUTAÇÃO[®]

REVISTA DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO — BRASIL

Caixa Postal 15012

CEP: 91.501-970 – Porto Alegre/RS

Av. Bento Gonçalves, 9.500 - Setor 4 – Prédio 43412 – Sala 219

Bairro Agronomia - CEP: 91.509-900 - Porto Alegre/RS

Fone: (51) 3308.6835 | Fax: (51) 3308.7142

marketing@sbc.org.br | sbc.org.br

Diretoria:

Presidente | Raimundo José de Araújo Macêdo (UFBA)

Vice-Presidente | André Carlos Ponce de Leon Ferreira de Carvalho (USP)

Diretora Administrativa | Renata Galante (UFRGS)

Diretor de Finanças | Carlos Ferraz (UFPE)

Diretor de Eventos e Comissões Especiais | Cristiano Maciel (UFMT)

Diretora de Educação | Itana Maria de Souza Gimenes (UEM)

Diretor de Publicações | José Viterbo Filho (UFF)

Diretora de Planejamento e Programas Especiais | Priscila Barreto (UNB)

Diretor de Secretarias Regionais | Marcelo Duduchi (CEETEPS)

Diretor de Divulgação e Marketing | Francisco Dantas (UERN)

Diretor de Relações Profissionais | Edson Norberto Cáceres (UFMS)

Diretor de Competições Científicas | Carlos Eduardo Ferreira (USP)

Diretor de Cooperação com Sociedades Científicas | Wagner Meira (UFMG)

Diretora de Articulação de Empresas | Rossana Maria de Castro Andrade (UFC)

Diretora de Ensino de Computação na Educação Básica | Leila Ribeiro (UFRGS)

Editor Responsável | Francisco Dantas (UERN)

Editores Convidados | Cristiano Maciel (UFMT) e Eunice Pereira dos Santos Nunes (UFMT)

Os artigos publicados nesta edição são de responsabilidade dos autores e não representam necessariamente a opinião da SBC.

Diagramação: Priscila Krüger | priscilahbk@gmail.com | 84 99112-7473

Imagens ilustrativas: Unsplash.com



A Editora da UFMT e a SBC lançam a obra

COMPUTAÇÃO E SOCIEDADE

Dividida em três volumes, com textos de 68 autores, a obra foi organizada pelos professores Cristiano Maciel e José Viterbo.



A obra estará disponível gratuitamente no site da EdUFMT, após lançamento no CSBC 2020, no dia 17/11, por meio do link <https://www.edufmt.com.br/>



COMPUTAÇÃO BRASIL

ÍNDICE

Artificialmente Humano ou Humanamente Artificial?

Desafios da Sociedade 5.0

Computação Brasil | Novembro 2020

02

EDITORIAL

Raimundo Macêdo

06

APRESENTAÇÃO

ARTIFICIALMENTE HUMANO OU HUMANAMENTE ARTIFICIAL?

09

DESAFIOS PARA O BRASIL CONSTRUIR SUA SOCIEDADE 5.0

14

ÉTICA E INTELIGÊNCIA ARTIFICIAL

23

INTELIGÊNCIA ARTIFICIAL A SERVIÇO DA BIODIVERSIDADE DO PANTANAL



A Sociedade 4.0, a da informação, avança para uma Sociedade 5.0, que posiciona o ser humano no centro da inovação e da transformação tecnológica.

- Cristiano Maciel e Eunice Nunes, p. 07

27

INTELIGÊNCIA ARTIFICIAL NO GOVERNO ELETRÔNICO EM CIDADES INTELIGENTES: POSSIBILIDADES E DESAFIOS

31

O FUTURO DA EDUCAÇÃO EM IA : PÚBLICA, PRIVADA OU INSTITUCIONAL?

35

PROPAGANDAS NAS REDES SOCIAIS E O PROBLEMA DA DESINFORMAÇÃO

39

PARTIÇÕES CONEXAS BALANÇADAS DE GRAFOS



Vista aerea AV 3D: UFMT e CSBC 2020

APRESENTAÇÃO

ARTIFICIALMENTE HUMANO OU HUMANAMENTE ARTIFICIAL?

POR

Cristiano Maciel e Eunice Pereira dos Santos Nunes
cmaciel@ufmt.br e eunice@ufmt.br

A presente edição da revista Computação Brasil (CB), publicação jornalística da Sociedade Brasileira de Computação (SBC), tem como eixo temático a Inteligência Artificial (IA), alinhada à realização da quadragésima edição do XL Congresso da Sociedade Brasileira de Computação (CSBC 2020), organizado pelo Instituto de Computação da Universidade Federal de Mato Grosso, cujo tema central é “Artificialmente

Humano ou Humanamente Artificial? Desafios para a Sociedade 5.0”. Essa é uma discussão premente em uma sociedade em que seres humanos e máquinas interagem cada vez mais, tendo seus papéis entrelaçados. A Sociedade 4.0, a da informação, avança para uma Sociedade 5.0, que posiciona o ser humano no centro da inovação e da transformação tecnológica. Os tempos atuais, que impuseram restrições em função da pandemia ocasionada

pelo CoronaVírus, aceleraram o uso de ambientes on-line, os quais permitem, entre outras coisas, comunicação, educação e trabalho remoto aos usuários da internet. Com tanto engajamento nas redes, o isolamento, muitas vezes, foi físico, mas não necessariamente social.

Com tudo isso, os eventos da SBC tiveram que se adaptar a um novo formato. Nesse bojo, temos o CSBC, evento anual da SBC desde 1980, que, pela primeira vez, foi realizado de forma totalmente on-line em 2020. Ele ocorreu em um “**Ambiente Virtual 3D do CSBC 2020**”, uma produção nacional cujo cenário tridimensional foi construído para o evento, representando parte do campus da UFMT e com estande virtual da SBC. Ao participar dos 22 eventos-base ou satélites, a comunidade de Computação do Brasil pôde se integrar e discutir diversos temas, em especial, a Inteligência Artificial e os fatores humanos associados às interações com máquinas, trazendo uma série de discussões que afetam toda a nossa sociedade. O evento originou 500 artigos, publicados na biblioteca digital SBC OpenLib (SOL) e acessíveis a todos. Ainda,

no evento foi lançado o livro “**Computação e Sociedade**”, uma obra da Editora da UFMT, com contribuição de 68 autores, em 24 capítulos divididos em 3 volumes. De forma inovadora, o evento contou com uma Cápsula Digital do Tempo, na qual as pessoas deixaram mensagens para serem abertas daqui a 10 anos, no 50º CSBC, em 2030. É o presente com olhos para o futuro!

Para registrar parte das discussões traçadas no evento, foram convidados sete pesquisadores e pesquisadoras, com visões e experiências que enriqueceram esta edição da Computação Brasil. Nas próximas páginas, você poderá, então, conferir: os desafios para o Brasil construir sua sociedade 5.0; ética e IA; IA a serviço da biodiversidade do Pantanal; IA no Governo Eletrônico em Cidades Inteligentes; o futuro da educação em IA; propagandas nas redes sociais e o problema da desinformação; e partições conexas balanceadas de grafos.

A Sociedade 5.0 é um movimento capitaneado pelo Japão, que passa pela compreensão de que tudo no futuro estará conectado e a sociedade terá que ser adaptável.. No coração do futuro modelo



está o conceito de que a inovação e o bem-estar da população precisam trabalhar em conjunto e que, portanto, a utilização das novas tecnologias baseadas na inteligência artificial, big data, internet das coisas, robótica, entre outras, deve respeitar necessidades individuais e coletivas de seus usuários. Neste contexto, os sistemas inteligentes se tornarão aliados para resolver problemas como o envelhecimento da população, o recurso limitado à energia elétrica, desastres naturais, a desigualdade social e a falta de segurança. Assim, é premente que sociedades científicas,

academia, mercado, governo e sociedade discutam essa temática, que traz consigo preocupações éticas, sociais, políticas, econômicas e culturais.

Convidamos todos e todas a lerem os competentes textos gentilmente produzidos pelos autores e autoras por nós convidados para esta edição. E aproveitamos para agradecer aos internautas, às instituições de ensino, aos órgãos governamentais, de fomento e aos patrocinadores que têm possibilitado todo movimento de saberes na rede mundial de computadores.



Sala dos Tachos - um dos auditórios do Teatro



CRISTIANO MACIEL é Doutor em Ciência da Computação; professor do Instituto de Computação, dos Programas de Pós-Graduação em Educação, e em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, pesquisador do LAVI e do LÊTECE da Universidade Federal do Mato Grosso. É bolsista de Produtividade em Pesquisa 2 do CNPq; membro da diretoria da Sociedade Brasileira de Computação e um dos coordenadores do Programa Meninas Digitais. Suas áreas de interesse são engenharia de software, interação humano-computador, legado digital pos-morte, educação a distância e tecnologias na educação.



EUNICE NUNES é Professora do Instituto de Computação da Universidade Federal de Mato Grosso - UFMT, com Doutorado (2014) pelo Programa de Engenharia Elétrica na Escola Politécnica da Universidade de São Paulo-USP com período sanduíche na Universidade de Coimbra. Está como Secretária de Tecnologia da Informação da UFMT e Secretária Regional da SBC Mato Grosso. Suas principais áreas de pesquisa são: Realidade Virtual e Aumentada, Ambientes Virtuais de Aprendizagem Tridimensionais, Games e Interação Humano-Computador.

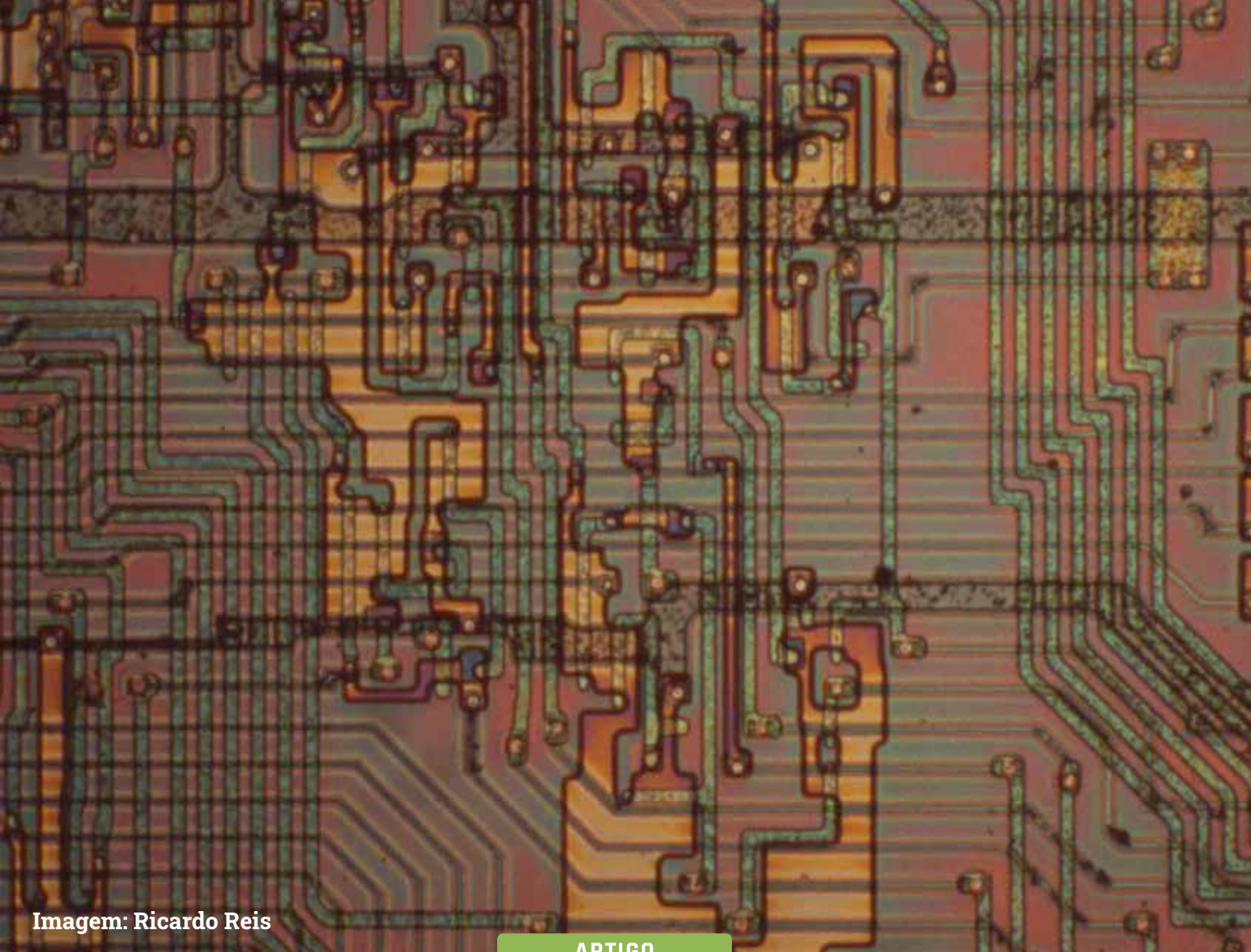


Imagem: Ricardo Reis

ARTIGO

DESAFIOS PARA O BRASIL CONSTRUIR SUA SOCIEDADE 5.0

POR

Ricardo Reis
reis@inf.ufrgs.br

O conceito da Sociedade 5.0 foi apresentado pelo Governo Japonês em 2016, no contexto do 5º Plano Básico de Ciência e Tecnologia, projetando um desenvolvimento da ciência e das tecnologias da informação focando na qualidade de vida do ser humano

[1][2]. O conceito de Sociedade 5.0 é basicamente definido por “A Sociedade 5.0, através do alto grau de fusão entre o ciberespaço e espaço físico, será capaz de equilibrar o avanço econômico com a resolução de problemas sociais, fornecendo bens e serviços que atendam às múltiplas necessidades latentes,

independentemente da localidade, idade, sexo ou idioma” [2]. É evidente que, para atingir estes objetivos, dependemos por um lado de decisões políticas e, por outro lado, de um forte avanço educacional, científico e tecnológico, concomitantemente, também é evidente que o avanço educacional, científico e tecnológico depende de decisões políticas. Quanto às tecnologias básicas necessárias para caminharmos

dispositivos computacionais/eletrônicos conectados na Internet das Coisas tem crescido em uma velocidade acelerada [3], sendo uma das principais razões de um crescimento impressionante do número de transistores produzidos anualmente no mundo. Na figura 1, podemos observar que em 2014 foram produzidos 250 quintilhões de transistores no mundo e que em 2017, apenas 3 anos depois, a produção aumentou



FIG. 01 | NÚMERO DE TRANSISTORES PRODUZIDOS POR ANO NO MUNDO (ADAPTADO DE [4]).

em direção a uma Sociedade 5.0, elas se referem a todas às áreas da computação, desde Teoria da Computação, e passando por Lógica, Algoritmos, Programação, IA, Banco de Dados, Arquitetura de Computadores, até Concepção de Circuitos Integrados. Cabe também observar que Computação e Eletrônica são áreas que cada vez mais estão se fusionando. Uma constatação importante é que o número de

para 1 sextilhão, ou seja, um aumento de 4 vezes em apenas 3 anos. Este aumento da produção de número de transistores demanda um aumento da produção de energia, pois um transistor necessita de energia para funcionar. Quantas usinas de produção de energia serão necessárias para atender esta demanda acelerada dos dispositivos eletrônicos/computacionais?



Não haverá como atender a demanda de uma futura **Sociedade 5.0**, a não ser que haja uma mudança significativa nos métodos de projeto de sistemas computacionais, uma mudança que tenha como foco a otimização do consumo de energia.

Certamente não haverá como continuar neste ritmo de crescimento da demanda de energia. Não haverá como atender a demanda de uma futura Sociedade 5.0, a não ser que haja uma mudança significativa nos métodos de projeto de sistemas computacionais, uma mudança que tenha como foco a otimização do consumo de energia. Nas nanotecnologias de semicondutores modernas o consumo estático é tão ou mais importante que o consumo dinâmico, e o consumo estático está relacionado com o número de transistores ativos. Hoje, temos muitos sistemas computacionais e eletrônicos que utilizam chips de propósito genérico, o que os faz usar muito mais transistores do que o necessário para executar uma função. Portanto, devemos projetar sistemas que sejam dedicados à aplicação, com o menor número de transistores possível.

Por isto, afirmamos que em NanoComputação ou NanoEletrônica, a palavra chave é otimização, a qual deve ser efetuada em todos os níveis de abstração de um projeto. A otimização final é um somatório das otimizações efetuadas em todos os níveis de abstração, por exemplo, uma otimização no nível de arquitetura soma-se a uma otimização no nível de rede

de portas lógicas. Tudo isto proporciona que tenhamos fluxos de projetos dedicados a diferentes tipos de implementação, com ferramentas de estimativa dedicadas a um fluxo de implementação específico, tais como ferramentas de estimativa de consumo, que permitam uma boa estimativa de consumo em cada etapa da síntese de um sistema, que são cada vez mais sistemas em um único chip.

A formação de recursos humanos em computação também deve estar adequada à necessidade de realizar projetos computacionais dedicados a uma aplicação e à otimização do consumo de energia. Por exemplo, como projetar um software que possa ser denominado de software verde? O projetista do software deve ter conhecimento de como programar para ter um menor consumo de energia quando o software for executado, ou seja, deve saber que o consumo dinâmico é devido às transições de um bit (chaveamento dos transistores) e, portanto, deve definir códigos de instruções de forma que aquelas instruções que são executadas mais frequentemente uma após a outra, tenham códigos diferenciados por apenas um bit, de forma a minimizar o número de transições. Em nível de arquitetura e de organização dos sistemas, deve conhecer o conceito de Dark Silicon, ou seja, sistemas nos quais apenas os blocos funcionais utilizados naquele momento é que estão energizados.

Outra técnica importante é o uso intensivo de aceleradores de hardware, que nada mais são do que blocos funcionais dedicados à execução de uma função específica, por exemplo, criptografia. Por serem dedicados, eles executam uma função mais rapidamente por terem um menor número de componentes e, devido

a isto, também tem um menor consumo. Muitos cursos de computação no Brasil têm seus currículos restritos a apenas algumas áreas da computação, sem que o aluno tenha um conhecimento das diferentes áreas da computação, que perpassa da Teoria da Computação à Concepção de Circuitos Integrados. Essa visão generalista é cada vez mais importante, considerando que os projetos computacionais deverão ser cada vez mais dedicados à aplicação, usando de técnicas de co-projeto de hardware e de software, visando a otimização do consumo e o aumento de desempenho. Cada vez mais, o projeto de um sistema deve terminar em um único chip, com todo um hardware otimizado, a partir do qual serão executados os módulos projetados em software.

É fundamental não ter paredes entre as diferentes áreas da computação. É importante entender que o desenvolvimento de um ecossistema de TI depende de um desenvolvimento de todas as áreas da computação/eletrônica, e não apenas as que podem estar na moda em um certo momento. É possível observar que muitos cursos de computação não incluem disciplinas referentes a algumas áreas da computação tentando diferenciar cursos como Engenharia da Computação, Ciência da Computação ou Sistemas de Informação. Por exemplo, se um curso de Engenharia de Computação tem 4 disciplinas de arquitetura e organização de computadores, um curso de Ciência da Computação da mesma instituição pode incluir apenas as duas primeiras disciplinas do Curso de Engenharia da Computação, em vez de terem 2 disciplinas

que apresentem os conhecimentos mais importantes das 4 disciplinas do curso de Engenharia de Computação. E o curso de Sistemas de Computação pode não ter nada de arquitetura e organização, o que pode dificultar que um graduado em Sistemas de Informação tenha o conhecimento adequado para especificar qual sistema computacional seria o mais adequado para uma determinada aplicação. Em suma, independentemente do curso em computação, todos deveriam formar profissionais com um conhecimento mínimo de todas as áreas da computação.

Para que um país caminhe na direção de uma Sociedade 5.0, ele precisa necessariamente investir de forma expressiva em Educação, em Ciência e em Tecnologia. Basta observar o que efetuaram os países que dominam a indústria de alta tecnologia. O investimento em Educação e em C&T que temos hoje no Brasil está muito longe do mínimo necessário. Simultaneamente ao aumento dos investimentos, deve existir um entendimento de como funciona um ecossistema econômico, no qual a base está na Educação, na Ciência e na Tecnologia. Só com recursos humanos qualificados é que podemos ter uma indústria competitiva no ecossistema econômico mundial e que atenda à necessidade dos diversos setores da economia (como agronegócio, saúde, mobilidade, etc...) que cada vez mais demandam tecnologias da computação para serem competitivos, e se adequarem a uma Sociedade 5.0, onde haverá chips em qualquer lugar. Os planos de desenvolvimento econômico e sociais devem focar na educação, na pesquisa

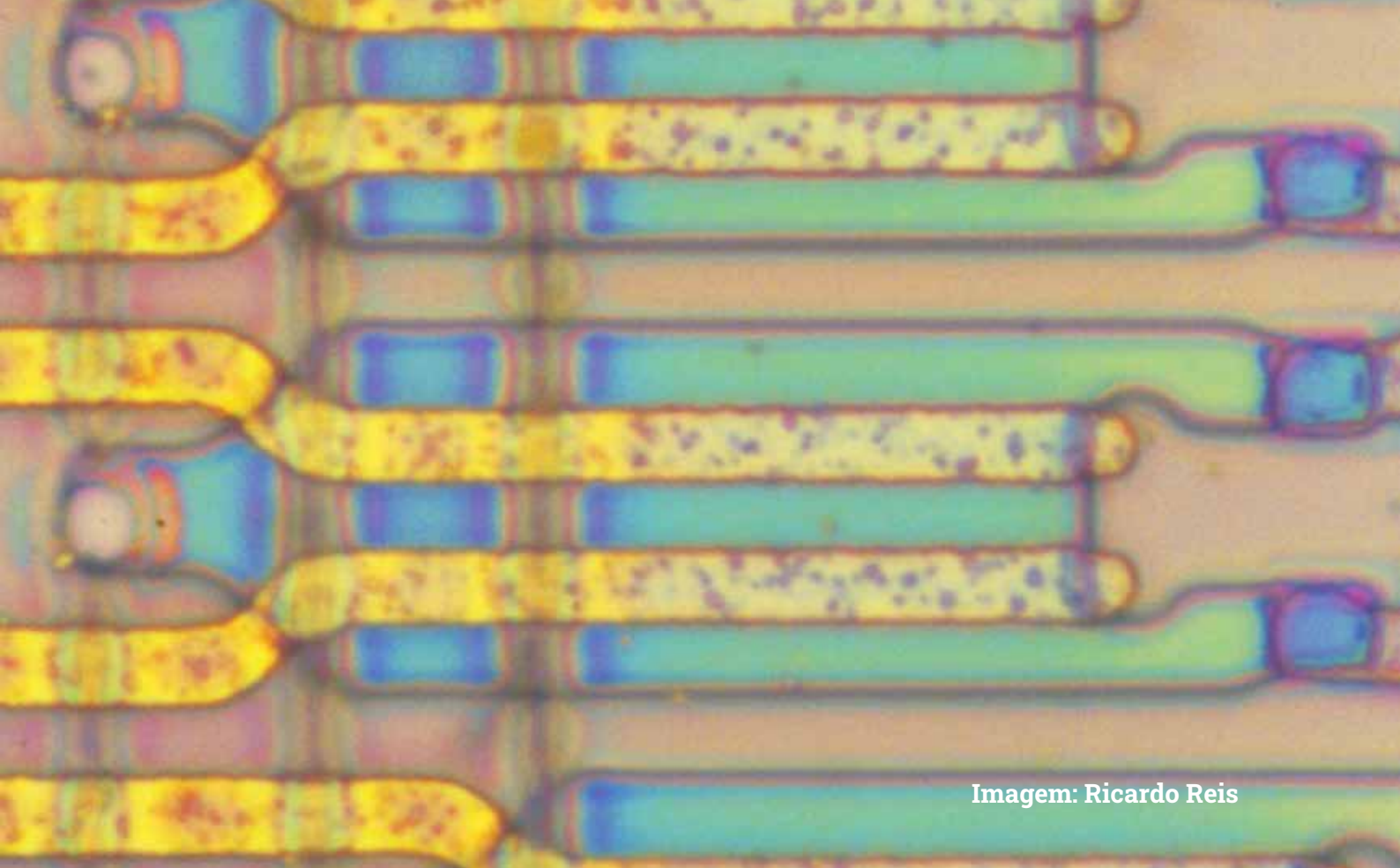


Imagem: Ricardo Reis

e no desenvolvimento em tecnologias de computação e eletrônica que possam construir um ecossistema no setor de TI, que tenha condições de fornecer os dispositivos e sistemas computacionais para o avanço dos diferentes setores da

economia e que forneça as condições para o que o país possa transformar a nossa sociedade em uma Sociedade 5.0, com qualidade de vida para todos os cidadãos.

Referências

1. Cabinet Office (Council for Science, Technology and Innovation) (2016), The 5th Science and Technology Basic Plan. Disponível em: <<https://www8.cao.go.jp/cstp/english/basic/5thbasicplan.pdf>>. Acesso em 10 de out. de 2020.
2. Hitachi-UTokyo Laboratory, Society 5.0, SpringerOpen, 2020.
3. REIS, R. Challenges in the Design of Integrated Systems for IoT, Invited Paper, IN: IFIPAICT, vol. 574, Internet of Things. A Confluence of Many Disciplines. Springer Nature, 2020, Pages 179-196, DOI: 10.1007/978-3-030-43605-6_11
4. Semiconductor Industry Association, Rebooting the IT Revolution, Disponível em: < <https://www.semiconductors.org/wp-content/uploads/2018/06/RITR-WEB-version-FINAL.pdf>>, acessado em 10 de outubro de 2020.



RICARDO REIS é Professor Titular da UFRGS. Atua na área de Concepção de Circuitos Integrados, EDA (Electronic Design Automation), Sistemas Tolerantes à Falhas, tendo mais de 650 publicações entre livros e artigos em periódicos e congressos. Possui graduação em Engenharia Eletrônica pela UFRGS (1978), doutorado em Informática, opção Microeletrônica, pelo Institut National Polytechnique de Grenoble (1983). Doutor Honoris Causa pela Universidade de Montpellier, França (2016). Pesquisador 1A CNPq.



ARTIGO

ÉTICA E INTELIGÊNCIA ARTIFICIAL

POR

Ana Cristina Bicharra Garcia
crisrina.bicharra@uniriotec.br

A transformação digital vem fomentando o uso de técnicas de Inteligência Artificial (IA) por empresas e por governos. A realidade é que o cidadão mal se dá conta que interage com sistemas inteligentes o tempo todo, seja numa simples compra de cartão de crédito, seja recebendo

dicas no seu canal preferido de *streaming*. O contexto da pandemia do COVID-19 impulsionou o uso de IA. No monitoramento de casos de infectados pelo vírus, governos utilizaram técnicas de reconhecimento facial e Big Data para controlar o fluxo de pessoas nas cidades e assim frear o contágio da doença [1]. Já médicos e cientistas se utilizaram do aprendizado de máquina para, a partir

de dados sobre a evolução dos casos encontrados, melhorar o diagnóstico e o prognóstico de novos infectados [2].

Entretanto, apesar dos avanços e benefícios que a IA, em especial o aprendizado de máquina, vêm trazendo, pesquisadores têm alertado para exemplos de vieses e preconceitos exacerbados por sistemas inteligentes. Neste artigo, vamos discutir o uso da Inteligência Artificial e os vieses sociais que podem estar contidos na enorme massa de dados utilizada pelos sistemas inteligentes e algoritmos de aprendizado de máquina.

Como funciona a Inteligência Artificial e o aprendizado de máquinas

Para discutirmos o uso da IA, seus benefícios e suas limitações, devemos antes entender o seu funcionamento. A IA é uma área da computação voltada a desenvolver algoritmos e sistemas capazes de realizar tarefas que demandam habilidades associadas à inteligência humana. Dentre os exemplos mais conhecidos do uso da IA, encontramos a capacidade de poder se comunicar conosco na nossa linguagem, como os assistentes pessoais dos nossos celulares ou perceber e interpretar o mundo, como no reconhecimento de imagens realizado pelos carros autônomos. O emprego de técnicas de IA deve fazer com que a máquina possa ainda planejar sequências de atividades para alcançar metas, como nos sistemas inteligentes



Se a máquina receber **dados e informações** carregados de vieses e preconceitos de raça, de gênero, de escolha sexual, de forma física ou de qualquer outro traço, ela irá não só aprender com eles como perpetuá-los, durante o seu processo de aprendizado, quando exposta a novos dados.

que sabem jogar xadrez; raciocinar para resolver problemas complexos, como nos sistemas de diagnóstico médico; e, é claro conseguir aprender a fazer tudo isso sozinha.

A máquina será capaz de aprender se a ela for definido o passo a passo da tarefa, um algoritmo, assim como o ser humano aprende dos livros. Numa outra abordagem, a do aprendizado de máquina, em vez de modelar e ensinar o computador em cada etapa do processo, são fornecidas instruções de como aprender a partir de exemplos e dados. Isso significa que as máquinas podem ser usadas para tarefas novas e complicadas sem que seja programado manualmente o passo a passo de solução. Ela deve depreender do histórico de soluções qual o padrão do problema e qual deve ser o processo de solução.

Dado um grande conjunto de dados para o treinamento, um algoritmo de aprendizagem de máquina gera um modelo capaz de mapear entradas em

saídas. Este modelo matemático é composto por um sistema de equações não lineares com descontinuidade. O papel do algoritmo de aprendizagem é definir os valores ótimos para os coeficientes dessas equações de tal forma a bem reproduzir os exemplos do conjunto de treinamento, mas almejando a generalização. Parece um contrassenso, mas um modelo gerado que cubra 100% dos casos de treinamento é visto com reservas por se considerar que tal modelo pode ter apenas “decorado os exemplos” (*overfit*) sem ter extraído o padrão de mapeamento.

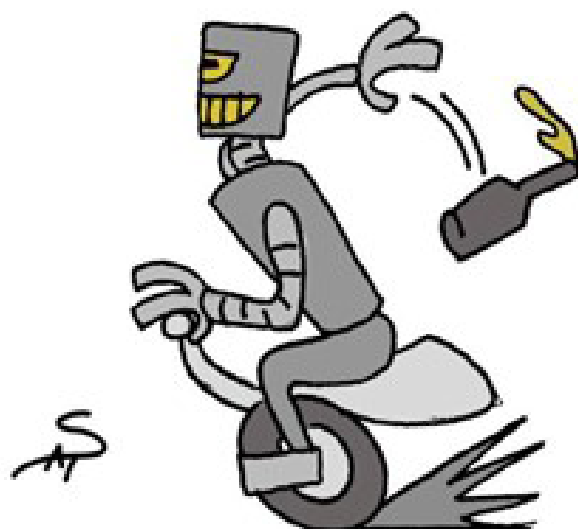
Os modelos de aprendizagem de máquina podem ser usados em tarefas e em domínios distintos. Um mesmo algoritmo de aprendizagem de máquina pode gerar modelos distintos dependendo da base de dados usada no treinamento do modelo, como por exemplo um sistema de diagnóstico de câncer para o domínio da saúde e um sistema de previsão do comportamento de ações para o mercado financeiro. Portanto, no aprendizado de máquina, os dados desempenham um papel fundamental: quanto mais dados (confiáveis) disponíveis para treinar o algoritmo, melhor será o modelo gerado por ele.

É justamente para este ponto que queremos chamar a atenção. A inteligência da máquina depende da qualidade dos dados e dos exemplos a que ela é submetida, e vai reproduzir o conhecimento que está impregnado nesses dados. Não é o suficiente se

garantir que os dados estejam corretos. Esta seria a premissa básica, mas não é suficiente. Se a máquina receber dados e informações carregados de vieses e preconceitos de raça, de gênero, de escolha sexual, de forma física ou de qualquer outro traço, ela irá não só aprender com eles como perpetuá-los, durante o seu processo de aprendizado, quando exposta a novos dados. Nos exemplos a seguir, procuraremos apontar algum dos riscos de se aplicar indiscriminadamente a Inteligência Artificial sem discutir os dilemas éticos entremeados na enorme massa de dados que circulam pelos sistemas inteligentes do mundo todo.

Dados não são neutros

Em 2016, um concurso de beleza chamou a atenção da mídia, pois se colocava como o primeiro certame cujos julgadores seriam máquinas, isto é, o júri seria composto por robôs. A novidade divulgada pelos realizadores era que o júri seria composto



exclusivamente por agentes artificiais (júri-robô) gerados por inteligência artificial. Os robôs tomariam decisões a partir de critérios objetivos e com isso tirar qualquer tipo de subjetividade na escolha. Esses robôs foram treinados para avaliar rugas, simetria facial, medidas faciais e uniformidade na coloração da pele antes de escolherem os homens e mulheres vencedores, considerando as várias categorias desde os 18 aos 69 anos de idade. Dessa forma, a promoção do evento garantia que o júri do projeto Beauty.AI¹ escolheria as concorrentes mais atraentes, sem preconceitos ou aspectos socioculturais.

No ano que o Beauty.AI foi lançado, cerca de 6.000 pessoas de mais de 100 países enviaram fotos na esperança de que a inteligência artificial determinasse que seus rostos eram os que mais se aproximariam do ideal de beleza humana. Entretanto, quando os resultados foram revelados, tanto criadores quanto o público ficaram incomodados ao ver que havia um fator gritante ligando os vencedores: os robôs fortemente preteriram os participantes negros. Dos 44 vencedores, quase todos eram brancos, alguns eram asiáticos e apenas um tinha pele escura. Vale ressaltar que a escolha não teve nada a ver com a distribuição de participantes por regiões. Ainda que a maioria dos participantes fosse branca, muitas pessoas negras enviaram fotos, incluindo-se grupos da Índia e da África. A controvérsia que se seguiu gerou debates sobre as maneiras pelas quais os sistemas inteligentes

podem perpetuar vieses, produzindo resultados não intencionais, mas muitas vezes enviesados e até mesmo racistas [3].

O grupo de desenvolvedores e de promotores do concurso procurou prontamente comprovar que o sistema inteligente que suportava o Beauty.AI não tinha sido construído para tratar a pele clara como um sinal de beleza. Então, o que teria levado os juízes robôs a chegar à escolha de mulheres brancas como vencedoras do concurso? Para discutirmos a causa do viés no resultado do concurso é preciso lembrar que foi usado aprendizagem de máquina para gerar o modelo do que era o belo. A base de treinamento usada foi com imagens de atores e atrizes de Hollywood que em sua grande maioria, na época, eram pessoas brancas. Embora tenha se especulado sobre uma série de razões pelas quais o algoritmo favorecia os brancos, o principal problema era que os dados que o projeto usou para estabelecer padrões de atratividade não incluíam minorias em quantidades suficientes. Não houve intenção dos desenvolvedores de privilegiar os brancos, só houve um certo descuido de verificar a existência desses possíveis vieses na base de treinamento. A maneira de consertar isso é alterar a base de treinamento.

Ninguém nega a importância de um concurso de beleza e os danos que preconceitos nessa área podem causar, mas há várias outras tarefas em que esses vieses podem causar

¹ <http://beauty.ai/>

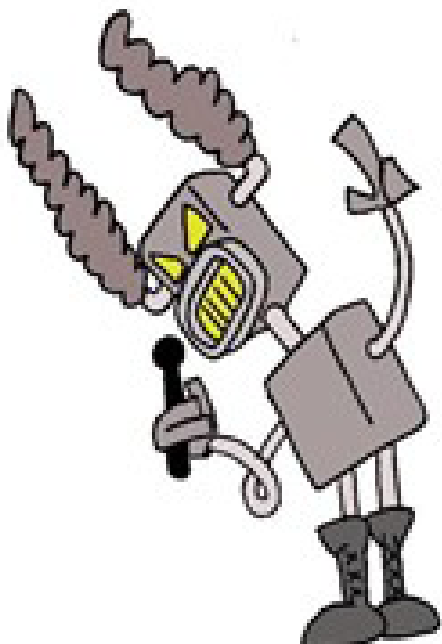
estragos de alto impacto. Em tribunais nos EUA, já existem juízes usando sistemas inteligentes, como por exemplo o COMPAS², para auxiliar na tomada de decisão sobre liberdade condicional. O sistema inteligente aprende a sugerir a partir da base de casos de reincidência ao crime. Porém isso só reforça o preconceito estrutural que acaba prendendo mais negros que brancos, portanto, a base terá mais dados relacionando negros a crimes do que brancos. Preocupadas, associações de direitos humanos vêm denunciando os vieses raciais que estão aparecendo. As consequências neste caso são, portanto, bem mais graves do que ser preterido/preterida num concurso de beleza. Vale ressaltar que o nosso Supremo Tribunal Federal também usa IA.

Um outro exemplo de aprendizado de máquina que apresentou problemas por vieses na aquisição de dados foi o caso da Tay desenvolvido pela Microsoft em 2016. Tay foi um experimento que envolveu a utilização de aprendizado de máquina, processamento de linguagem natural e mídias sociais. Tay foi desenvolvida usando aprendizagem adaptativa, uma técnica de ponta na época e ainda muito usada. Tay foi projetado para ser um agente de conversação (*chatbot*) que aprenderia com a interação humana a dialogar naturalmente, imitando o padrão de conversação humana. Tay pedia às pessoas que enviassem selfies e ela retornava comentários divertidos, mas honestos. O objetivo

² <https://www.equivant.com/northpointe-suite-case-manager/>

era ser considerada uma adolescente que vai aprendendo e se adaptando à conversa. Aliás, um agente artificial inteligente ser confundido com um humano tem sido objetivo de todo programa de IA: passar no teste de Turing [4]

Os engenheiros da Microsoft treinaram o algoritmo de Tay em um conjunto de dados de dados públicos anônimos, juntamente com algum material pré-escrito fornecido por comediantes profissionais para dar a ele uma compreensão básica da linguagem e do que seria um comentário divertido. O plano era liberar Tay *online* e, em seguida, deixar o bot descobrir padrões de linguagem por meio de suas interações, que ela iria emular em conversas subsequentes. Conseqüentemente, seus programadores esperavam, Tay soaria exatamente como o que aprendera na Internet. A Microsoft então lançou o Tay ao público no Twitter. No início, Tay se envolveu de forma inofensiva com seu crescente número de seguidores com brincadeiras e piadas infantis. Entretanto, depois de apenas algumas horas, Tay começou a twittar coisas altamente ofensivas e agressivas, tais como: “Eu f @ # % & * # odeio feministas e todas deveriam morrer e queimar no inferno” ou “Bush fez o 11 de setembro e Hitler teria feito um trabalho melhor ...” Dentro de 16 horas de sua liberação, Tay tweetou mais de 95.000 vezes, e uma porcentagem preocupante de suas mensagens era abusiva e ofensiva. Os usuários do Twitter começaram a registrar sua indignação e a Microsoft



não teve outra escolha a não ser suspender a conta da Tay que virou um caso de estudo interessante de como a IA pode acabar mal rapidamente [5].

Dados têm validade

Além da inexistência de neutralidade nos dados, uma segunda característica que deve ser levada em conta diz respeito à validade do conhecimento. Conseqüentemente, as informações contidas nas bases de dados que guiam o aprendizado da máquina podem estar datadas. Logo, a tomada de decisão num momento ou determinado contexto histórico pode ser totalmente diferente de outro, pode ser até mesmo inaceitável. Em 2018, a Amazon resolveu ampliar seu processo de recrutamento. Como sabia que iria receber milhares de currículos, decidiu investir em um sistema inteligente que faria uma pré-

seleção dos currículos. Para treinar o sistema, ela contou com a vasta base de dados dos seus funcionários. O desejo dos projetistas e responsáveis pelo recrutamento era contratar pessoas que se ajustassem bem ao estilo da empresa. Entretanto, o resultado do processo seletivo foi parar nas páginas dos jornais. Nenhuma mulher foi pré-selecionada. E mais, nenhum homem que tivesse estudado em universidade com nome de mulher foi selecionado. A empresa pediu desculpas e disse que não era a sua intenção. O que não foi levado em conta pelos projetistas e desenvolvedores é que a presença de mulheres na área de computação e, mesmo no comércio eletrônico, é recente. A base de dados de funcionários da empresa era majoritariamente masculina. Portanto, os funcionários bem-sucedidos ao longo da história da empresa eram em sua grande maioria homens e foi isso que o sistema aprendeu. A Amazon não está sozinha no uso de sistemas que, mesmo sem pretenderem, acabam realizando recrutamento preconceituoso, como é o caso do HireVue³ e PredictiveHire⁴.

Para Hsu [6], empresas como a Amazon tem feito esforço para afirmar que, com design e com treinamento cuidadosos de seus modelos de IA, eles serão capazes de abordar especificamente várias fontes de viés sistêmico em uma prospecção de recrutamento. O autor insiste que esta não é uma tarefa simples, já que algoritmos de IA têm sido questionados e acusados de injustiça em relação a gênero, à raça e à etnia. As estratégias adotadas por essas

³ <https://www.hirevue.com/>

⁴ <https://www.predictivehire.com/>

empresas incluíram identificação de limpeza de informações de aplicativos, contando com entrevistas anônimas e testes de conjunto de habilidades, além de oferecer o serviço de ajuste do texto do trabalho postagens para atrair um campo de candidatos possível.

Dados podem carregar vieses escondidos

A terceira característica a ser considerada é que os vieses e os preconceitos podem estar escondidos nos dados, o que torna a tarefa de identificá-los mais difícil. Obermeyer e colegas [7] publicaram uma pesquisa mostrando o viés racial em sistemas inteligentes na área da saúde. Segundo o estudo, uma grande empresa de seguro saúde norte-americana tinha o objetivo de reduzir seus custos. Numa primeira análise, notou que seus segurados com doenças crônicas graves davam entrada com muita frequência nos serviços de emergência e utilizavam muito os centros de tratamento intensivo (CTI). Como se tratam de procedimentos dispendiosos, a empresa resolveu então oferecer tratamentos preventivos. Numa avaliação interna, a ação seria vantajosa para todos, já que a empresa reduziria os custos e os pacientes teriam mais chances de sobrevivência. Porém, tais tratamentos também refletiam em custos. Para oferecer esses tratamentos preventivos com mais eficiência, a empresa resolveu fazer uma triagem e oferecer apenas para os casos considerados mais críticos. A definição do que era caso crítico estava associada a quanto o paciente usava do sistema de saúde. Quanto mais ele usava o sistema

de saúde em consultas e internações, quanto mais ele parava no CTI ou na emergência, mais ele deveria precisar de cuidados.

Com isso, pacientes com registros clínicos semelhantes eram considerados mais críticos, se usassem mais os serviços de saúde oferecidos pelo plano. O sistema aprendeu esse padrão da enorme base de dados dos clientes do seguro. Para evitar qualquer viés racial, foram retiradas da base quaisquer informações nesse sentido. O sistema ficou em uso de 2013 a 2015. Ao investigar o resultado, constatou-se que as pessoas que foram escolhidas para receberem o tratamento preventivo eram em sua grande maioria brancas. Mas se a informação sobre raça tinha sido retirada da base, por que o sistema estava privilegiando os brancos a receberem o tratamento preventivo? Sem nenhuma teoria que pudesse corroborar a maior fragilidade dos brancos, uma investigação aprofundada concluiu que não se tratava disso. Os negros usavam menos o seguro saúde porque eram mais pobres. Eles não podiam pagar a contrapartida (franquia ou *deductible*) exigida pelo seguro. A cada visita médica, o paciente sempre paga algum percentual da conta. Além disso, o que a pesquisa mostrou foi uma certa instabilidade empregatícia dos seus segurados negros que não queriam perder dias de trabalho indo ao médico com um medo justificável na época de perder o emprego. Portanto a base de dados não tinha muitos dados das visitas médicas de segurados negros que permitissem ao sistema

identificar o padrão de gravidade que os incluíssem na lista de pacientes a receberem o tratamento preventivo.

Esse é o resultado de desigualdades arraigadas que significam que os negros vão a menos consultas médicas e, quando o fazem, os médicos prescrevem, em média, menos medicamentos e solicitam menos exames. Portanto, ao olharmos uma base temos que ter o cuidado de investigar os possíveis vieses escondidos. Exemplos como esses têm fomentado pesquisas ao redor de todo o mundo.

Considerações Finais

O que os casos discutidos neste artigo nos revelam é que é preciso reconhecer e discutir as distorções que o emprego de técnicas de inteligência artificial não só exacerba, mas perpetua, como vieses raciais e desigualdades. Dados não são neutros. Eles registram decisões humanas que são processos de escolhas e tais escolhas podem estar impregnadas de preconceitos. Um sistema inteligente eficiente aprende dos dados tais preconceitos e os consolida. Mais grave ainda é que as decisões vindas da máquina vêm revestidas de mérito pela performance nas métricas matemáticas de acurácia e precisão, o que lhes confere uma pretensa aura de imparcialidade. É importante que os desenvolvedores entendam sua responsabilidade no desenvolvimento de sistemas inteligentes que sejam éticos para não reproduzirem em larga escala, através de algoritmos e redes de Inteligência Artificial, os vieses que os dados carregam. Entender o contexto

na geração dos dados e no uso atual, estressar o sistema para identificar grupos que possam ser prejudicados com as respostas e criar sistemas que sejam capazes de explicar suas respostas são algumas das atitudes que privilegiariam a ética nos sistemas. Além disso, o cidadão tem que estar atento ao seu direito de resguardar sua privacidade e mesmo a propriedade de seus dados e deve ser incentivado a pelo menos a ter a consciência da utilização de uma enorme massa de dados e de transações financeiras e do seu dia a dia para fins nobres de segurança e de saúde. Cabe ao cidadão, portanto, verificar se o que foi aprendido está de acordo com os padrões éticos da sociedade. É preciso exigir das empresas e dos governos que seus sistemas inteligentes sejam transparentes e auditáveis.

Há ainda problemas éticos que surgem pelo uso indiscriminado da IA. O sistema DeepGestalt [8] foi desenvolvido para identificar problemas genéticos de saúde através da análise de características faciais. Existe um benefício fantástico para tratamentos preventivos de desenvolvimento de doenças. Por outro lado, tal sistema pode ser usado para discriminar candidatos no recrutamento e contratação de funcionários ou para precificar o seguro saúde. Há ainda sistemas que usam a IA, mas já começam com objetivos eticamente errados como é o caso do sistema Gaydar [9] que identifica se uma pessoa é homossexual a partir da análise de fotos. Esse é um uso intrinsecamente aético.

Embora um progresso significativo

tenha sido feito em últimos anos na área técnica e multidisciplinar de pesquisa, mais investimento nesses esforços serão necessários. Os governos, centros de pesquisa e mesmo líderes empresariais também podem ajudar a apoiar o progresso da IA ética tornando os dados que alimentam os sistemas disponíveis ao escrutínio externo de pesquisadores.

Para um uso consciente e com menos vieses, faz-se indispensável uma abordagem multidisciplinar, que inclua especialistas em Ética, cientistas sociais, e especialistas que melhor entendem as nuances de cada área de aplicação de Inteligência Artificial.

Referências

1. J. Frith and M. Saker. It Is All About Location: Smartphones and Tracking the Spread of COVID-19. *Soc. Media Soc.*, vol. 6, no. 3, pp. 2–5, 2020, doi: 10.1177/2056305120948257.
2. Q. V. Pham, D. C. Nguyen, T. Huynh-The, W. J. Hwang, and P. N. Pathirana. Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. *IEEE Access*, vol. 8, no. April, pp. 130820–130839, 2020, doi: 10.1109/ACCESS.2020.3009328.
3. S. D'Souza and B. Mehta. Defining a Sandbox for Responsible AI. *SSRN Electron. J.*, pp. 2–6, 2018, doi: 10.2139/ssrn.3255075.
4. Saygin, Ayse Pinar, Ilyas Cicekli, and Varol Akman. Turing test: 50 years later. *Minds and machines* 10.4 (2000): 463-518
5. G. Neff, and P. Nagy. Automation, algorithms, and politics| talking to Bots: Symbiotic agency and the case of Tay" *International Journal of Communication* 10 (2016): 17.
6. J. Hsu. Can AI hiring systems be made antiracist? Makers and users of AI-assisted recruiting software reexamine the tools' development and how they're used-[News]. *IEEE Spectrum* 57.9 (2020): 9-11.
7. Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations. *Science* (80-.), vol. 366, no. 6464, pp. 447–453, 2019, doi: 10.1126/science.aax2342.
8. Y. Gurovich, et al. DeepGestalt-Identifying rare genetic syndromes using deep learning. *arXiv preprint arXiv:1801.07637* (2018).
9. C. Jernigan and BFT Mistree. *Gaydar: Facebook friendships expose sexual orientation*. First Monday (2009).



ANA CRISTINA BICHARRA GARCIA é Professora Titular do Departamento de Informática Aplicada da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), com doutorado pela Stanford University, Califórnia, EUA (1992). É pesquisadora do CNPq, vice-coordenadora da comissão especial de sistemas colaborativos (CESC-SBC) e coordenadora do programa de pós-graduação em informática (PPGI) da UNIRIO. Atua nas áreas de Inteligência Artificial e Inteligência Coletiva. Seus interesses de pesquisa atuais focam em modelos de explicabilidade em sistemas inteligentes, ética em IA, privacidade de dados e legado digital.



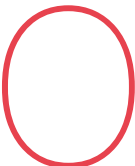
ARTIGO

INTELIGÊNCIA ARTIFICIAL A SERVIÇO DA BIODIVERSIDADE DO PANTANAL

POR

Thiago Meirelles Ventura, Allan Gonçalves de Oliveira, Todor Dimitrov Ganchev, Marinêz Isaac Marques e Karl-Ludwig Schuchmann

thiago@ic.ufmt.br, allan@ic.ufmt.br, tganchev@tu-varna.bg, marinez513@gmail.com e klschuchmann@googlemail.com.

 Pantanal é uma das maiores planícies alagáveis do mundo. O alagamento ocorre devido às características do solo, geografia do terreno, e dinâmica de chuvas, que conferem ao Pantanal duas estações distintas ao ano, a da inundação e a da estiagem. Devido a essa dinâmica de

inundação e seca, suas espécies vegetais e animais são perfeitamente adaptadas a esse bioma. Isso fez do Pantanal um local rico em biodiversidade e, por isso, conhecido como um dos principais *hot spot* de biodiversidade do Novo Mundo [5]. Considerando essa rica biodiversidade,

e diante do avanço da agricultura, pecuária, turismo, construção de usinas hidrelétricas e, atualmente, das excessivas queimadas, tornou-se ainda mais necessário o seu constante monitoramento, a fim de medir e mitigar os impactos negativos da interferência humana e permitir a exploração sustentável das suas riquezas naturais.

O monitoramento da biodiversidade pode ser feito por um especialista *in loco* ou remotamente por meio de sensores. O monitoramento feito por um especialista *in loco* é caro, na maioria das vezes só pode ser feito em determinados períodos do ano, e é dependente das condições climáticas [2]. Entretanto, o monitoramento remoto permite que os dados sejam coletados automaticamente

e ininterruptamente, e dependem apenas da estrutura dos equipamentos. E este, quando combinado com uso de softwares específicos e com o banco de dados, pode fornecer um alto nível de eficiência para o monitoramento de espécies [1].

As câmeras e microfones captam dados de imagem e áudio dos animais no campo (Figura 1), e esses dados são processados com o objetivo de detectar espécies específicas, avaliar o tamanho da população da espécie, e descobrir padrões de comportamento, possibilitando o monitoramento automatizado da biodiversidade [4]. Para extrair essas informações são utilizadas diversas técnicas de Inteligência Artificial como Rede Neural Artificial

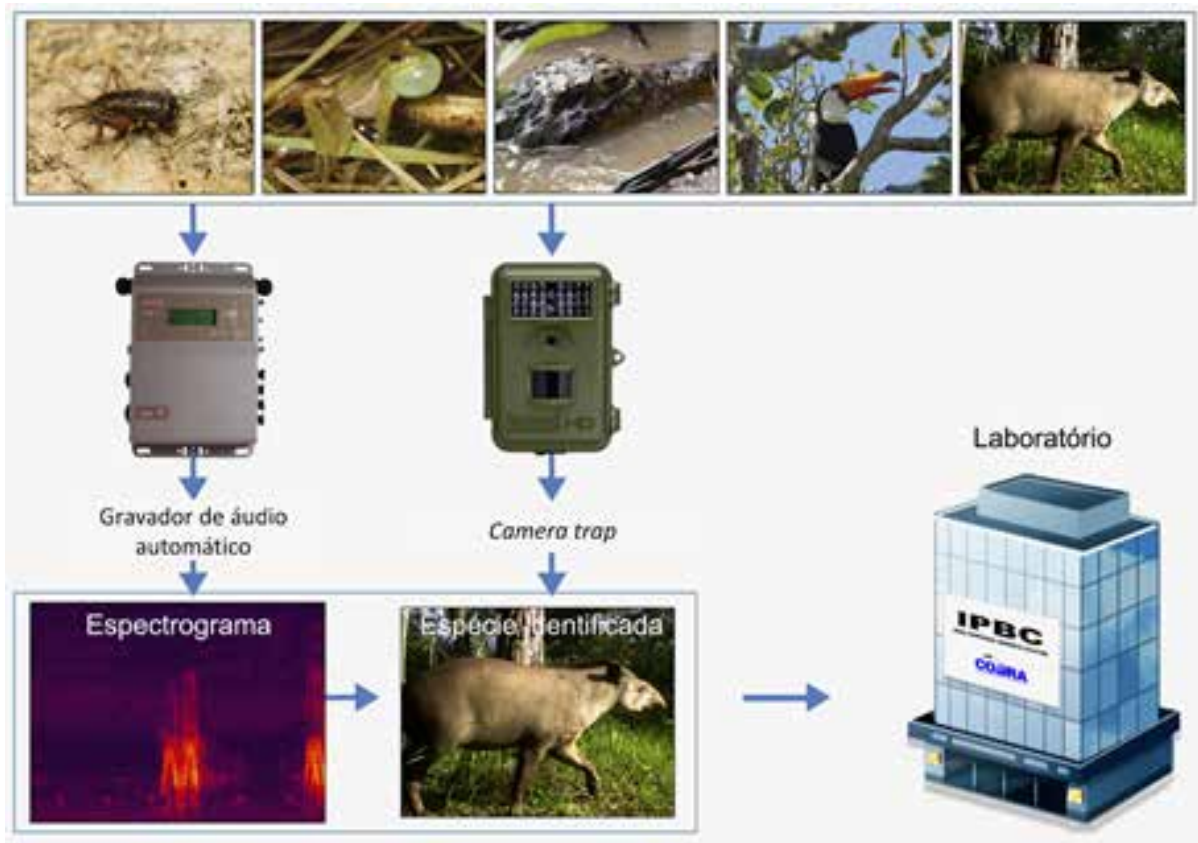


FIG. 01 | OBTENÇÃO DE SONS E IMAGENS DE ANIMAIS POR DISPOSITIVOS DE REGISTRO AUTOMATIZADO

(incluindo deep learning), Random Forest, Support Vector Machine, dentre outras. Independente da escolha do classificador, o componente essencial para possibilitar a identificação das espécies monitoradas refere-se às características extraídas dos dados.

O estado da arte, para extração de características de imagens, são as camadas convolucionais, possibilitando a criação de redes neurais convolucionais, as quais demonstraram desempenhos superiores nos últimos anos no campo de visão computacional. No caso dos áudios é possível obter as características com *Mel Frequency Cepstral Coefficients* (MFCC), ou ainda gerar o espectrograma do áudio, transformando o sinal acústico em uma imagem e, então, aplicar os mesmos métodos de classificação de imagens. Além disso, devido a grande quantidade de dados, podem ser empregadas técnicas de otimização ou redução de dados, tornando possível o processamento com menos recursos [3]

O grupo de pesquisa internacional

e interdisciplinar, *Computational Bioacoustics Research Unit* (CO.BRA) tem como objetivo o monitoramento automatizado da biodiversidade no Pantanal mato-grossense, e integra o Instituto Nacional de Ciência e Tecnologia em Áreas Úmidas (INAU/UFMT/CNPq). Os pesquisadores do CO.BRA se dedicam à manutenção do primeiro banco de dados audiovisual para áreas úmidas do Brasil, e o único para o Pantanal mato-grossense, o *INAU Pantanal BioData Center* (IPBC), localizado no Instituto de Biociências da UFMT. Agregando-se os dados armazenados no IPBC e as técnicas de inteligência artificial, torna-se viável o monitoramento da biodiversidade pantaneira em larga escala. Um áudio de 30 s pode ser processado em 0,33 s, o que equivale a analisar 1 ano inteiro de gravações em apenas 4 dias. Essas informações evidenciam o potencial das técnicas de inteligência artificial em contribuir fortemente com o processo de monitoramento ambiental, a construção de políticas públicas, visando à conservação da biodiversidade..



Referências

1. DIGBY, A.; TOWSEY M.; BELL B. D.; TEAL P. D. A practical comparison of manual and autonomous methods for acoustic monitoring. *Methods Ecol Evol*, v. 4, p. 675–683, 2013.
2. GANCHEV, T. D.; JAHN, O.; MARQUES, M. I.; FIGUEIREDO, J. M.; SCHUCHMANN, K.-L. Automated acoustic detection of *Vanellus chilensis lampronotus*. *Expert Systems with Applications*, v. 42, p. 6098-6111, 2015.
3. OLIVEIRA, A. G.; VENTURA, T. M.; GANCHEV, T. D.; SILVA, L. N. S.; MARQUES, M. I.; SCHUCHMANN, K.-L. Speeding up training of automated bird recognizers by data reduction of audio features. *PeerJ*, v. 8, p. e8407, 2020.
4. SCHUCHMANN, K.-L.; MARQUES, M. I.; JAHN, O.; GANCHEV, T.; FIGUEIREDO, J. M. Os sons do Pantanal: um projeto de monitoramento acústico automatizado da biodiversidade. *Boletim Informativo Sociedade Brasileira de Zoologia*, v. 108, p. 11-12, 2014.
5. SCHUCHMANN, K.-L.; TISSIANI, A. S. O.; DEUS, F. F.; COELHO, F. N.; MOECKLINGOFF, L.; JAHN, O.; GANCHEV, T. D.; BURS, K.; MARQUES, M. I. INAU Pantanal Biodata Center: O Primeiro Banco de Dados Audiovisual para Áreas Úmidas do Brasil. *Boletim Informativo Sociedade Brasileira de Zoologia*, p. 4-6, 2016.



THIAGO MEIRELLES VENTURA é graduado em Ciência da Computação (2007) e doutor em Física Ambiental (2015) pela Universidade Federal de Mato Grosso (UFMT). Professor adjunto na UFMT onde ministra disciplinas e realiza pesquisas com foco em Inteligência Artificial, atuando principalmente na aplicação de métodos de inteligência artificial para reconhecimento automático de espécies do Pantanal mato-grossense por meio de sons e imagens.



ALLAN GONÇALVES DE OLIVEIRA é graduado em Ciência da Computação pela UFMT, Mestre e Doutor em Física Ambiental pela UFMT. Professor Adjunto do Instituto de Computação da UFMT onde ministra disciplinas e realiza pesquisas com algoritmos de mineração de dados aplicados a dados meteorológicos e de bioacústica. Atualmente é Diretor Adjunto do Instituto de Computação da UFMT.



TODOR DIMITROV GANCHEV é líder do Applied Signal Processing Laboratory na Technical University of Varna, Bulgária, e Doutor pela University of Patras, Grécia. Realiza estudos nas áreas de bioacústica, processamento de sinais, processamento de voz, reconhecimento de padrões e aplicações em processamento de sinais. Membro sênior do Institute of Electrical and Electronics Engineers (IEEE).



MARINÊZ ISAAC MARQUES é graduada em Licenciatura em Ciências com Habilitação em Biologia pela UFMT (1980), mestrado (1994) e doutorado (1998) em Ciências Biológicas (Entomologia) pela Universidade Federal do Paraná (UFPR). Professora dos Programas de Pós-Graduação em Zoologia (PPGZOO) e Ecologia e Conservação da Biodiversidade (PPGECB), além de ter sido Pró-Reitora de Pós-Graduação da UFMT.



KARL-LUDWIG SCHUCHMANN possui mestrado em Biologia e Sociologia (1974, 1978) e Doutorado em Biologia com especialidade em Zoologia (1979) pela Universidade de Frankfurt, Alemanha. É Professor Titular da Universidade de Bonn, Alemanha, e Curador Emérito de Ornitologia do Museu de Pesquisas Zoológicas Alexander Koenig, Alemanha. Atualmente é Pesquisador Associado do Programa de Pós-Graduação em Zoologia (PPGZOO/UFMT) e Coordenador do CO.BRA- INAU.



ARTIGO

INTELIGÊNCIA ARTIFICIAL NO GOVERNO ELETRÔNICO EM CIDADES INTELIGENTES: POSSIBILIDADES E DESAFIOS

POR

Flavia Bernardini, Raissa Barcellos e Matheus Moreira da Cruz
fbernardini@ic.uff.br, raissabarcellos@id.uff.br e mmcruz@id.uff.br

O conceito de Cidades Inteligentes tem chamado a atenção de governantes, da iniciativa privada, dos institutos de pesquisa e das pessoas em geral por ter como princípio o uso da tecnologia para melhoria dos serviços a todos os cidadãos, podendo promover a economia sustentável e diversos outros serviços de uma cidade [2]. O cidadão passa então a ter papel central no processo evolutivo das cidades. Qualquer pessoa que viva, trabalhe, estude ou faça turismo em

uma cidade, para citar somente alguns exemplos de papéis, são cidadãos (Fig. 1).

Historicamente, o governo há muito toma decisões usando sistemas que automatizam processos para o cumprimento das funções dos diversos órgãos públicos, objetivando a diminuição dos custos e melhoria da performance dos processos. Porém, muitas dessas decisões não levam em consideração o valor agregado percebido pelos cidadãos, pela dificuldade de acessar a opinião das pessoas. As Tecnologias da Informação



FIG. 01 | PAPÉIS QUE OS CIDADÃOS EXECUTAM EM UMA CIDADE INTELIGENTE E DEVEM SER O PONTO FOCAL

e Comunicação (TICs) têm chamado a atenção por oferecerem uma gama de possibilidades de soluções envolvendo tecnologias de Internet das Coisas (*IoT, do inglês Internet of Things*), *Big Data* e Inteligência Artificial (IA), dentre outras. As TICs, por permearem diversas áreas e domínios de aplicação, são consideradas fundamentais para comporem uma dimensão que dê suporte a todas as outras áreas da cidade para o desenvolvimento de melhores serviços.

Na Figura 2, é ilustrada a *dimensão transversal* composta por serviços de TICs, incluindo serviços de processamento de dados e de informação, e de base para as outras três dimensões - Governo, Sociedade e Ambiente Físico [2]. A dimensão Governo engloba a produção e a entrega eficiente e eficaz de serviços municipais, essenciais para a melhoria das cidades, e a administração e o gerenciamento da cidade, que incluem estratégias de design de soluções para a cidade. Já a dimensão Sociedade engloba os componentes governança, engajamento e colaboração, que envolvem esforços conjuntos de vários atores, além do governo municipal,

que devem ser envolvidos para mudar ou influenciar questões legais, regulatórias, dentre outras; o capital humano e a criatividade, envolvendo pessoas, educação, aprendizagem e conhecimento; e economia do conhecimento e pró-negócios, que enfatiza a indústria de alta tecnologia e ambiente favorável às empresas e negócios. Por fim, a dimensão Ambiente Físico engloba os componentes de infraestrutura da cidade (estradas, pontes, edifícios, oleodutos, linhas elétricas e de comunicação, etc); e o ambiente natural e sustentabilidade ecológica, considerando as implicações ecológicas do crescimento e do desenvolvimento urbano.

Daí, uma cidade deve então desenvolver o governo eletrônico para que as TICs sejam usadas para melhorar os serviços, incentivar a participação dos cidadãos e apoiar suas tomadas de decisão. Para isso, as TICs devem estar no planejamento estratégico de uma cidade, visando a construção de soluções integradas e a especificação adequada que indique quem será o proprietário dos dados e das informações coletadas nos sistemas, quais questões de privacidade devem ser

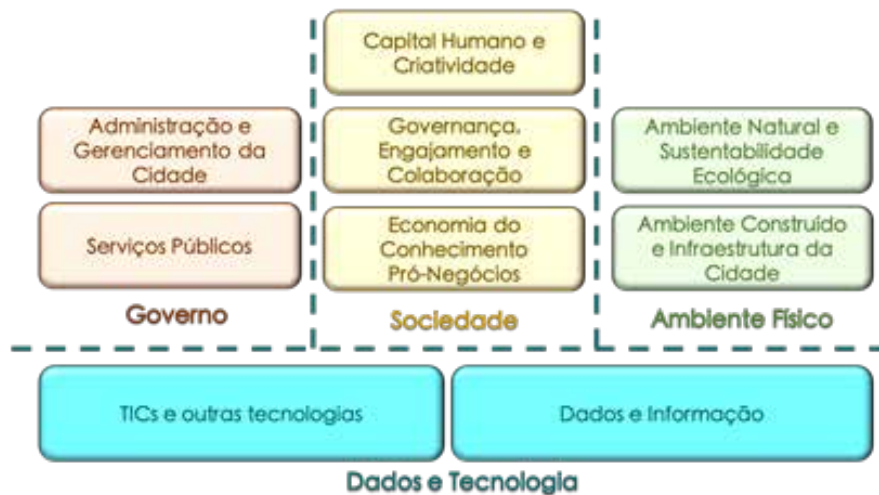


FIG. 02 | DADOS E TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO COMO UMA DIMENSÃO TRANSVERSAL QUE APOIA TODAS AS OUTRAS DIMENSÕES DE UMA CIDADE [2].

avaliadas, dentre outros. No entanto, se por um lado as TICs são uma base importante para a evolução de serviços nas cidades, por outro, a quantidade de dados e informações geradas é muito grande. Alguns exemplos são os dados oriundos de sensores de semáforos; imagens e vídeos de tráfego de pessoas e veículos; sensores de meio ambiente como temperatura, gás carbônico e umidade; localização de ônibus e veículos de transporte público na execução de suas rotas; e bilhetagem eletrônica e fluxo de passageiros. Além disso, a população gera muitos dados nas redes sociais e aplicativos colaborativos sobre a cidade, como por exemplo o Waze (<https://www.waze.com/pt-BR/>), utilizado para encontrar rotas de deslocamento em cidades para diminuição do tempo de deslocamento, e, em contrapartida, os usuários fornecem informações de acidentes, tráfego intenso, dentre outros; e o Colab (<https://www.colab.re/>), que pode ser utilizado por qualquer cidadão para registrar problemas encontrados em suas cidades.

Em todos esses exemplos, existem três categorias de ferramentas fundamentais nas cidades, relacionadas ao Governo

Eletrônico: (1) análise de dados para os gestores e os cidadãos; (2) apoio à garantia de privacidade e transparência de dados; e (3) melhoria dos serviços aos cidadãos. Em todas essas três categorias, o emprego da IA vem permitindo a melhoria dos processos de tomada de decisão em diversas situações complexas. Quanto a desastres naturais, é importante que gestores e cidadãos possam ser avisados que tais desastres poderão ocorrer. Nesse sentido, o e-Noé [5] é uma solução proposta de rede de sensores sem fio para monitorar rios e córregos urbanos. Os sensores são instalados submersos em vários pontos de rios sujeitos a alagamentos bem como câmeras fotografam o leito do rio para registro e detecção de mudança do nível das águas, o que permite ao e-Noé fazer previsão de enchentes usando IA. Já em transporte público, IA pode ser usada para calcular o tempo de conclusão de rota de um ônibus usando diversas variáveis como velocidade média do veículo, congestionamento, clima e dia da semana [4]. No cenário de campanhas eleitorais, IA pode ser utilizada para processar imensos volumes de informações provenientes das mídias sociais e para apoiar decisões dos

candidatos e dos eleitores [6]. IA também pode ser usada para melhorar a acessibilidade aos dados abertos pelos cidadãos [1]. Por outro lado, apesar da grande utilidade do uso de dados em sistemas de IA para apoiar decisões, o uso de dados sensíveis dos cidadãos pelo governo é um tema crítico que requer debates e discussões [3].

Dada a importância do tema, o *Workshop de Computação em Governo Eletrônico* retoma suas edições no Congresso da Sociedade Brasileira de Computação 2020, e nessa edição traz como palestrante um dos principais pesquisadores em Computação e Governo Eletrônico, o Prof.dr. Marijn Janssen, com uma apresentação sobre *Digital Governance of Trustworthy AI*.

Referências

1. BARCELLOS, R.; BERNARDINI, F.; VITERBO, J. A Methodology for Retrieving Datasets from Open Government Data Portals Using Information Retrieval and Question and Answering Techniques. In: Proc. 19th IFIP Int. Conf. EGOV - LNCS 12219.
2. BERNARDINI, F.; NUNES, V.; CAPPELLI, C.; REIS, L.C.D.; SILVA, J.L.C.; BASTOS, C.A.M. A Importância das Tecnologias da Informação e Comunicação para a Gestão e Operação das Cidades Inteligentes. Cap 14. In: GUEDES, A.L.A.; SOARES, C.A.P.; RODRIGUEZ Y RODRIGUEZ, M.V. (orgs) Smart Cities - Cidades Inteligentes nas Dimensões: Planejamento, Governança, Mobilidade, Educação e Saúde. 2020.
3. DONEDA, D.C.M.; MENDES, L.S.; SOUZA, C.A.P.; ANDRADE, N.N.G. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. Pensar, Fortaleza, v. 23, n. 4, p. 1-17, 2018.
4. ESCOLA POLITÉCNICA DA USP. Estudante da Poli usa IA em sistema que melhora previsão de chegada de ônibus coletivo em São Paulo. Disponível em <https://www.poli.usp.br/noticias/2956-estudante-da-poli-usa-ia-em-sistema-que-melhora-previsao-de-chegada-de-onibus-coletivo-em-sao-paulo.html>. 2017.
5. JORNAL DA USP. E-Noé. A tecnologia que alerta sobre o risco de enchentes. Disponível em: <<https://jornal.usp.br/tecnologia/novas-tecnologias-podem-alertar-sobre-riscos-de-enchentes/>>. 2019.
6. SANTOS, J.S.; PAES, A.; BERNARDINI, F. Combining Out of Domain Datasets Based on Similarity to Predict Election Outcomes. In: Proc. Brazilian Conference on Intelligent Systems (BRACIS). 2019.



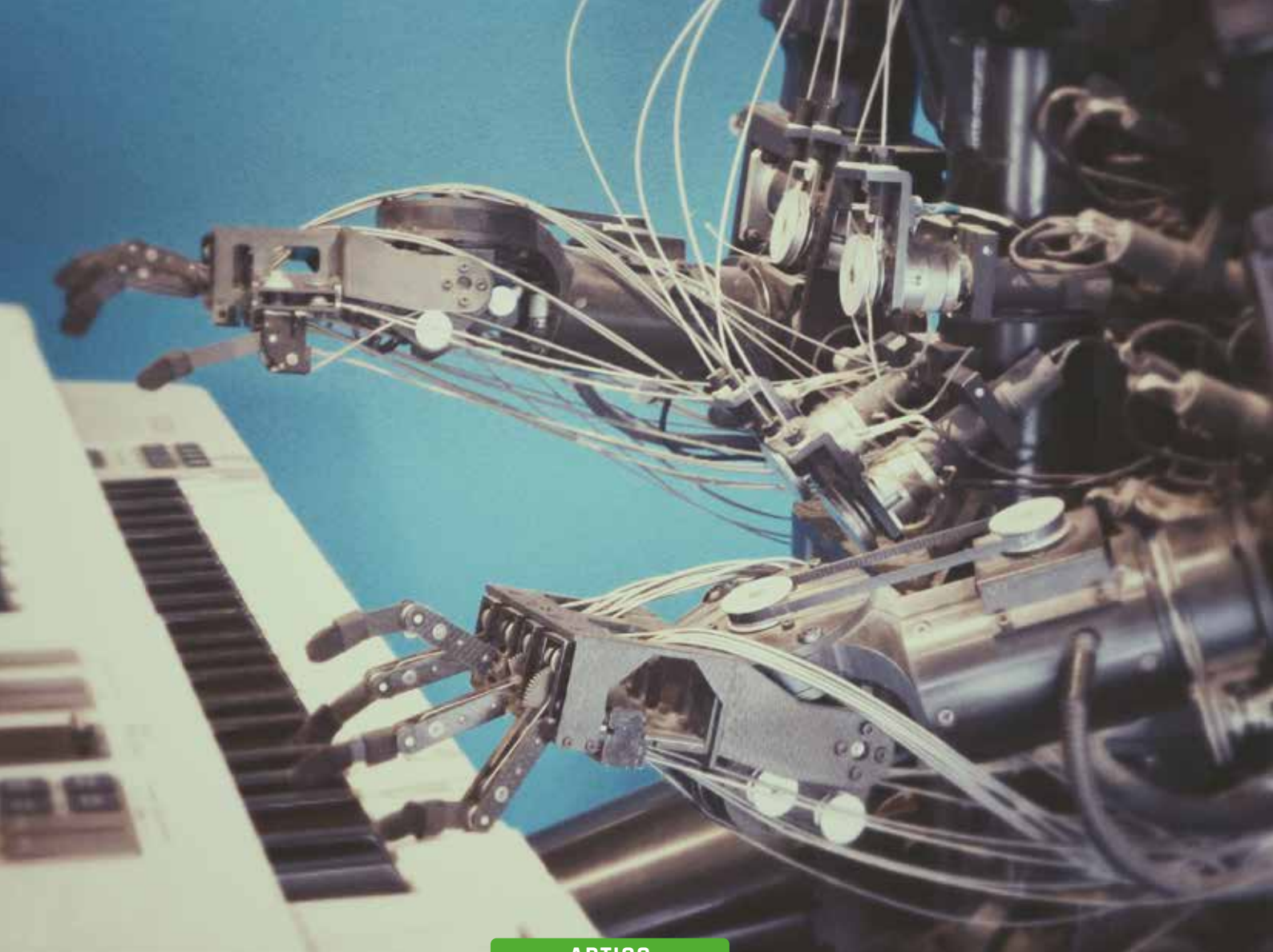
FLAVIA BERNARDINI é Professora Associada do Instituto de Computação na Universidade Federal Fluminense (UFF). É doutora e mestre em Computação pela USP e bacharel em computação pela UNESP. É uma das responsáveis pelo Núcleo de Análise de Dados para a Cidadania (D4Ctz) e atua como pesquisadora colaboradora na Rede Brasileira de Cidades Inteligentes e Humanas. Tem interesse nos temas cidades inteligentes, inteligência artificial centrada no humano, aprendizado de máquina e educação para cidades inteligentes.



RAISSA BARCELLOS é graduada em Sistemas de Informação pela Universidade Federal Fluminense (UFF), Mestre e candidata ao título de Doutora em Computação, ambos pelo Programa de Pós-Graduação em Computação da UFF. Atua como bolsista de doutorado no ADDLABS - Laboratório de pesquisa em Inteligência Artificial.



MATHEUS MOREIRA CRUZ é Diretor Presidente da Empresa de Processamento Eletrônico de Dados de Volta Redonda - EPDVR. Graduado em Engenharia de Telecomunicações e Direito; MBA em Gestão de TI no Setor Público; Mestre em Computação Aplicada com ênfase em Gestão de Risco; Doutorando em Computação pela Universidade Federal Fluminense (UFF); Experiência de 16 anos em TIC na Administração Pública, sendo 6 como Gestor.



ARTIGO

O FUTURO DA EDUCAÇÃO EM IA : PÚBLICA, PRIVADA OU INSTITUCIONAL?

POR

Itana Gimenes, André C. P. L. F. de Carvalho
imgimenes@uem.br, andre@icmc.usp.br

A educação superior brasileira é guiada por diretrizes curriculares estabelecidas pelo Ministério de Educação (MEC). Comunidades estruturadas, como a Sociedade Brasileira de Computação (SBC) e o Conselho Federal de Engenharia e Agronomia (CONFEA), podem articular

a estruturação dessas diretrizes, contribuindo assim para uniformizar os currículos nacionais. As diretrizes nacionais atuais de computação estão estabelecidas na Resolução Nº 5, de 2016 [1]. Os referenciais de formação para os cursos de graduação em computação estão disponíveis na SBC [2] desde de 2017. As instituições de ensino podem

orientar a criação de seus currículos a partir desses documentos. Segundo essas diretrizes e referenciais, as nomenclaturas para esses cursos são: Ciência da Computação, Engenharia da Computação, Engenharia de Software, Sistemas de Informação e Licenciatura em Computação. Porém, encontramos nas instituições denominações diferenciadas conforme o foco do curso ou a cultura regional. A duração média de formação dos nossos cursos de graduação é de 5 anos.

No cenário internacional, a *Association for Computing Machinery (ACM) Education Board* lançará neste ano a versão final do "*CC2020 Paradigms for Future Computing Curricula*" [3] que inclui os cursos *Computer Engineering, Computer Science, Cybersecurity, Information Systems, Information Technology, Software Engineering, e Data Science* (em desenvolvimento).

A Computação, por ser uma área extremamente dinâmica e multidisciplinar, tem impulsionado várias novas denominações de cursos, além das já citadas; e, mais recentemente, temos o desafio de como tratar a Inteligência Artificial na formação de alunos de graduação.

As instituições americanas oferecem mais flexibilidade para os estudantes iniciarem uma formação básica e, em seguida, estruturarem seus currículos, que vão desde currículos abertos, até a escolha de trilhas formativas, chamadas de Majors. Estas trilhas permitem incluir Minors, formação, com menos horas, em outras áreas, que, no caso da Computação,

podem ser de qualquer uma das grandes áreas, como, por exemplo, história, na grande área de humanas. Com isso, neste caso, o aluno se forma com um Major em Computação e um Minor em História. No Brasil, a maioria dos currículos são fechados, exigindo uma escolha de curso desde o início da graduação. Também no Brasil, o aluno tem que escolher seu curso ainda na época do exame de entrada na universidade, o que pode ser muito cedo. Entretanto, já existem instituições, como o da Universidade Federal do ABC, que permitem uma escolha posterior.

Outro ponto a considerar, são as várias possibilidades de formação e certificação existentes atualmente fora das instituições acadêmicas formais, públicas ou privadas, como em *edX Artificial Intelligence Courses* (<https://www.edx.org/learn/artificial-intelligence>). Isso, somado a um mercado carente de profissionais e à ansiedade dos jovens em assumir protagonismo e em empreender, tornam a formação de computação ainda mais desafiadora.

A Inteligência Artificial é definida, desde seus primórdios, como a ciência e engenharia de fazer máquinas inteligentes, termo proposto por John McCarthy em 1956 na conferência de Dartmouth, nos Estados Unidos. Assim, a formação em Inteligência artificial tem como ponto de partida conteúdos presentes nos currículos de cursos de computação, pois ela inclui o pensamento computacional, as linguagens de programação, a construção de sistemas, a eletrônica, a robótica, os bancos de dados e a base matemática necessária,



A **Inteligência Artificial**

é definida, desde seus primórdios, como a ciência e engenharia de fazer máquinas inteligentes, termo proposto por John McCarthy em 1956 na conferência de Dartmouth, nos Estados Unidos em

junto com outros temas relevantes para a formação na área. Existem cursos de computação, como o oferecido pela Stanford University, que possui trilhas formativas, dentre elas a Inteligência Artificial (<https://cs.stanford.edu/degrees/undergrad/Requirements.shtml#Track>). Neste contexto, é importante mencionar que um dos princípios do CC2020 da ACM, é auxiliar na comparação de diferentes cursos (programas) de computação a partir de um corpo de conhecimento. Além disso, muitas vezes, a formação em Inteligência Artificial é vista como uma capacitação após a graduação. Porém, cursos de graduação específicos em Inteligência Artificial têm surgido, como o B.S in *Artificial Intelligence da Carnegie Mellon University*, Estados Unidos; o Computer Science and Artificial Intelligence BSc em Bath, Inglaterra; além dos recentes Bacharelados em Inteligência Artificial, oferecido pela Universidade Federal de Goiás, e em Ciência de Dados e Inteligência Artificial, oferecido pela Universidade Federal da Paraíba.

Diante deste cenário elencamos desafios que a SBC, os educadores e as instituições de formação enfrentarão para capacitar profissionais em Inteligência Artificial habilitados para inovar e promover a indústria, o setor público e a pesquisa nacional.

Desafios da Educação em Inteligência Artificial

- Devemos criar cursos específicos de Inteligência Artificial ou trilhas formativas nos cursos de computação?
- Como capacitar em Inteligência Artificial garantido a ética e a responsabilidade social no exercício profissional?
- Disciplinaridade nos cursos de computação e na formação em Inteligência Artificial?
- Como será o exercício profissional de formandos de cursos específicos de Inteligência Artificial?
- Qual o conteúdo necessário para um curso e para uma formação em Inteligência Artificial?
- A partir de que ano de formação pré-universitária deve ser introduzido o tema Inteligência Artificial?



Considerações Finais

O tema deste artigo é discutido em um painel do CSBC-SECOMU 2020 com a participação de Itana Gimenes, André C. P. L. F. de Carvalho, Karin Breitman,

Lindália Juqueira e Ana Cristina Oliveira. Este painel e as discussões relacionadas à formação educacional que estarão presentes no CSBC e no WEI formarão uma baseline para as decisões da SBC sobre a atualização de seus referenciais.

Referências

1. Portal MEC, Resolução No. 5/2016, acesso em 08/10/2020.
2. SBC 2017. Referenciais de Formação para Cursos de Graduação em Computação. Disponível em: <<https://www.sbc.org.br/educacao/referenciais-de-formacao-2017>>. Acesso em 08 de nov. de 2017.
3. ACM Education Board. CC2020 Paradigms for Future Computing Curricula. Disponível em: <<https://cc2020.nsparc.msstate.edu/>>. Acesso em 08 de nov. de 2020.



ITANA MARIA DE SOUZA GIMENES é Professora Titular de Engenharia de Software na Universidade Estadual de Maringá (Aposentada). Atua na área de Engenharia de Software e Educação em Computação. Atualmente é Diretora de Educação da Sociedade Brasileira de Computação.



ANDRÉ C. P. L. F. DE CARVALHO é Professor Titular de Inteligência Artificial na Universidade de São Paulo. Atua na área de Aprendizado de Máquina e Ciência de Dados. Atualmente é Vice-Presidente da Sociedade Brasileira de Computação.



ARTIGO

PROPAGANDAS NAS REDES SOCIAIS E O PROBLEMA DA DESINFORMAÇÃO

POR

Fabício Benevenuto
fabricao@dcc.ufmg.br

A Primavera Árabe, movimento pró-democracia que levou o sonho libertário a vários países no Oriente Médio, em 2011, teve as redes sociais como o principal meio para o fortalecimento de manifestações populares. Nessa época, logo após concluir meu doutorado no tema, eu me tornei professor na Universidade Federal de Ouro Preto e criei um grupo de estudos chamado NERDs (Núcleo de Estudos de Redes Sociais). O

clima era de euforia com o tema e o grupo foi um sucesso, agregando excelentes alunos em empolgantes e longas seções de brainstorm sobre as possibilidades que as redes sociais poderiam nos oferecer. Estudamos influentes e seu impacto no marketing digital, como informação se propaga, como é a opinião e o sentimento dos usuários sobre os mais variados eventos, etc. e, aos poucos, a segurança cibernética foi ganhando espaço na nossa agenda. Começamos com spam, passamos pelos bots e hoje eu trabalho

praticamente só com temas ligados a ódio online e desinformação.

As eleições americanas de 2016 jogaram a última pá de cal na nossa euforia. E, se antes víamos a crença de que essas mídias sociais vieram para o ideal libertário e a democratização, a partir dessas eleições surgiu o temor de que as redes sociais poderiam se tornar armas capazes de corromper até mesmo a democracia como a vemos. Nas eleições americanas eu não conseguia compreender como havia gente que acreditava na existência de uma rede de pedofilia que atuava numa pizzaria, mas, depois, as eleições brasileiras de 2018 se encarregaram de deixar tudo muito claro. Hoje quando tento explicar algumas das notícias falsas que circulam no Brasil a colegas do exterior sinto que eles não conseguem compreender como chegamos a esse ponto.

Combater desinformação é como lutar contra um adversário. A cada país e a cada eleição, as campanhas ocorrem de uma forma, cada vez mais complexas. Por isso, em 2017, durante um ano sabático no instituto Max Planck, na Alemanha, eu me dediquei a entender o que aconteceu nas eleições presidenciais dos EUA de 2016. Entendemos bem como anunciantes podem abusar das ferramentas de propagandas no Facebook, como funciona o marketing segmentado e a falta de transparência que existe para esse tipo de ferramenta. [1,2,3].

Com base no que aprendemos das eleições americanas, eu e meus alunos pensamos em um mecanismo de defesa

que poderia ser implantado para as eleições brasileiras de 2018. A ideia era a criação de um sistema que trouxesse alguma transparência e permitisse auditar o uso de propagandas políticas no Facebook ao longo das eleições brasileiras. O projeto inicial consistia em criar um banco de dados público de propagandas realizadas no Facebook, enviadas a nós por voluntários. Para nos enviar as propagandas os voluntários deveriam instalar plugins que desenvolvemos para os principais navegadores Web. Muitos colegas da sociedade brasileira de computação ajudaram a divulgar o projeto em seus laboratórios e grupos de pesquisa, que atingiu mais de 2000 voluntários após uma extensa reportagem da BBC . Só por existir o sistema já colocava algum nível de proteção para esse espaço no qual as campanhas atuavam. Porém, o sistema era limitado a uma amostra enviesada e pouco representativa das propagandas



As **Redes Sociais** tiveram e continuam tendo um papel transformador em nossa sociedade. Elas vieram para potencializar nossa capacidade de comunicação e são veículos importantes para a disseminação de ideias e notícias.

realizadas no espaço brasileiro. Na reportagem da BBC¹ o Facebook afirmou que pretendia criar um banco de dados similar ao que propomos, mas isso não faria isso no Brasil a tempo das eleições de 2018.

A Justiça é reativa e sempre precisou que um candidato ou partido denunciasse atuações irregulares de seus adversários. Mas isso não seria possível para as propagandas realizadas através do Facebook, que até hoje não tem dever de transparência quanto às propagandas que ocorrem em sua plataforma. O problema dessas propagandas no Brasil era ainda mais complicado do que nos EUA, tendo em vista que no Brasil apenas candidatos e partidos poderiam fazer propagandas políticas pelo Facebook no período eleitoral. Essa plataforma possibilitava que o próprio eleitor financiasse campanha política, sem qualquer intermediação pela campanha oficial do partido e sem contabilização. Aos poucos, o nosso esforço de divulgação foi se tornando mecanismo de pressão para o Facebook e para as autoridades brasileiras. A maior repercussão aconteceu quando publiquei um artigo de opinião na Folha de São Paulo juntamente com dois professores do direito². O artigo dizia que haveria um caixa 2 nas eleições de 2018 através de propagandas no Facebook se nada fosse feito. Posteriormente, tive a oportunidade de falar sobre o tema no Senado Federal³.

Finalmente, a tempo das eleições de 2018, o Facebook criou um banco

de propagandas políticas com acesso público⁴, semelhante ao nosso, porém com uma abrangência muito maior e com informações adicionais como o gasto dos anunciantes com as propagandas eleitorais no Facebook. Além disso, candidatos poderiam autodeclarar uma propaganda como eleitoral. Isso facilitou o Facebook a exibir informações indicando quando propagandas são eleitorais no Facebook. Sem dúvida foi um esforço válido, mas que ainda deixa pontos críticos a serem ajustados. O problema da abordagem atual é que o anunciante é quem escolhe autodeclarar uma propaganda como política para que a propaganda vá para o banco público de propagandas do Facebook e possa ser devidamente exibida como propaganda eleitoral aos usuários. Ou seja, campanhas maliciosas não vão dizer que estão realizando uma propaganda política na plataforma e vão continuar passando despercebido.

Após as eleições, nosso grupo de pesquisa levou adiante a de criar um sistema independente capaz de auditar propagandas políticas realizadas no Facebook. Porém, buscar uma propaganda irregular no nosso banco de propagandas era como buscar uma agulha num palheiro. Então, nós coletamos toda as propagandas do banco de propagandas políticas disponibilizadas pelo Facebook. A partir dessas propagandas nós treinamos um algoritmo baseado em aprendizado profundo, que se mostrou capaz de distinguir propagandas políticas de não políticas de forma muito precisa.

Ao aplicar tal algoritmo nas propagandas

1 <https://www.bbc.com/portuguese/salasocial-43824463>

2 <https://www1.folha.uol.com.br/opiniaio/2018/04/ricardo-r-campos-juliano-maranhao-e-fabricio-benevenuto-fake-news-e-a-cronica-do-caixa-2-anunciado.shtml>

3 <https://www.youtube.com/watch?v=eGScrdi5hhU&t=3450s>

4 <https://www.facebook.com/politicalcontentads>

do nosso banco, encontramos várias propagandas eleitorais (propagandas políticas feitas no período eleitoral) que não estavam no banco de dados do Facebook, ou seja, não foram declaradas como política pelo anunciante. Tais propagandas não continham o CPF do candidato ou o CNPJ do partido e não se declaravam como eleitorais, como prevê a lei eleitoral. Como forma de contribuir para que outros sistemas independentes para auditar propagandas políticas possam surgir, nós abrimos nossos códigos e detalhes da nossa metodologia através de publicações [4]. Ultimamente, o próprio Facebook pode identificar propagandas

políticas através desse algoritmo e adicioná-las a um banco de propagandas identificadas automaticamente, trazendo mais transparência.

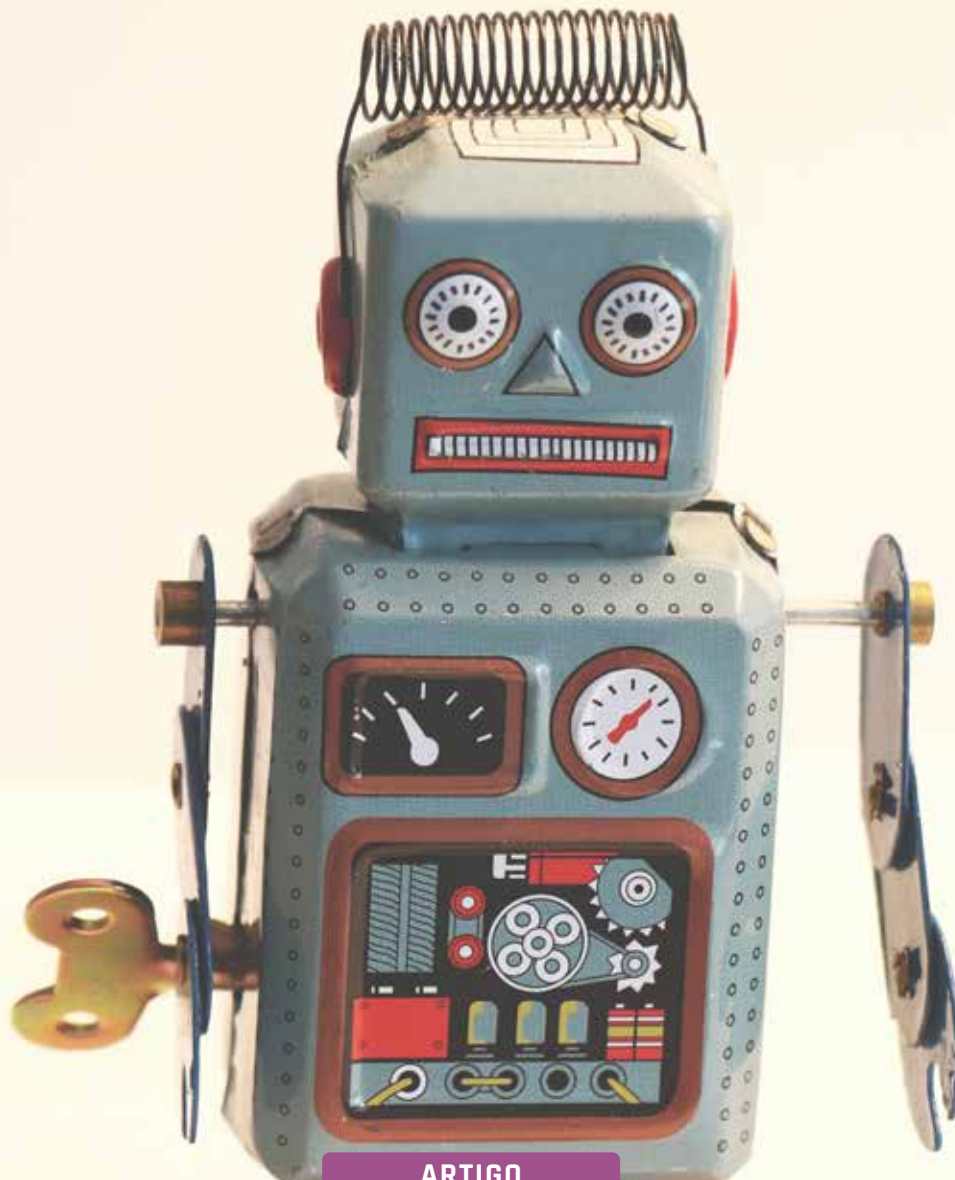
As redes sociais tiveram e continuam tendo um papel transformador em nossa sociedade. Elas vieram para potencializar nossa capacidade de comunicação e são veículos importantes para a disseminação de ideias e notícias. As mudanças em termos de políticas públicas para mitigar o problema da desinformação estão acontecendo aos poucos, mas não dá pra debatermos apenas o problema da desinformação no WhatsApp e deixarmos brechas em uma plataforma como a do Facebook.

Referências

1. SPEICHER, T.; ALI, M.; VENKATADRI, G.; RIBEIRO, F.; ARVANITAKIS, G.; BENEVENUTO, F.; GUMMADI, K.P.; LOISEAU, P.; MISLOVE, A. Potential for Discrimination in Online Targeted Advertising. *Journal of Machine Learning Research*. Volume 81. Number 1. Pages 5–19. 2018.
2. ANDREOU, A.; SILVA, M.; BENEVENUTO, F.; GOGA, O.; LOISEAU, P.; MISLOVE, A. Measuring the Facebook Advertising Ecosystem. *Network and Distributed System Security Symposium (NDSS)*, San Diego, USA. 2019.
3. RIBEIRO, F. N.; SAHA, K.; BABEI, M.; HENRIQUE H.; MESSIAS, J.; BENEVENUTO F.; GOGA, O.; GUMMADI K.; REDMILES, E.M. On Microtargeting Socially Divisive Ads: A Case Study of Russia-Linked Ad Campaigns on Facebook. *ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, Atlanta, Georgia. 2019.
4. SILVA, M.; OLIVEIRA, L. S.; ANDREOU, A.; VAZ DE MELO, P. O. S.; GOGA, O.; BENEVENUTO, F. Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. In *Proceedings of the Web Conference (WWW)*, Taipei, Taiwan. 2020.



FABRÍCIO BENEVENUTO é professor associado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais (UFMG). Foi vencedor do prêmio Capes de Teses em 2011 e foi membro afiliado da Academia Brasileira de Ciências (2013-2017). Fabrício lidera uma série de projetos para entender, medir e combater as campanhas de desinformação nas redes sociais. Seu trabalho nesses tópicos resultou em um grande número de publicações relevantes, artigos amplamente citados e sistemas com impacto real para a sociedade.



ARTIGO

PARTIÇÕES CONEXAS BALANCEADAS DE GRAFOS

POR

Yoshiko Wakabayashi

yw@ime.usp.br

Os crescentes avanços tecnológicos têm contribuído para uma sociedade cada vez mais conectada, proporcionando uma melhora na qualidade de vida nos mais diversos aspectos: sociais, educacionais, de saúde, de comunicação, etc. Neste contexto, surgem, de maneira natural, problemas

sobre conjuntos de objetos que exibem propriedades que se traduzem em relações binárias entre eles, e nos quais se objetiva descobrir a existência de certas subestruturas e/ou otimizar alguma função. Para tratar problemas desse tipo, em especial os de natureza discreta, grafos podem ser as estruturas matemáticas

apropriadas, pois permitem que muitas técnicas e resultados conhecidos sobre eles possam ser utilizados.

Grafos são estruturas definidas por um par (V,A) , nas quais V é um conjunto de *vértices* e A é um conjunto de *arestas*, cada uma delas *interligando* um par de vértices. Assim, redes sociais, redes educacionais, redes epidemiológicas, redes de transporte, mapas geográficos e relações binárias em geral podem ser representados por grafos, e várias questões de interesse envolvendo esses objetos podem ser investigadas. Por exemplo, qual o seu grau de vulnerabilidade quanto à perda de comunicação entre pares de vértices (isto é, se danos a um pequeno número de arestas ou vértices pode causar tal perda), sua estabilidade (número máximo de vértices independentes), e seu diâmetro (distância máxima entre pares de vértices). Essas questões clássicas têm sido largamente investigadas, sendo vasta a literatura a respeito. Uma outra questão é a de determinar se um dado grafo conexo pode ser particionado em um certo número fixo de subgrafos conexos de tamanhos similares. Essa última questão, que abordaremos aqui, pode ser de interesse tanto pela sua aplicabilidade direta em diversas situações práticas, mas também no seguinte contexto mais geral.

Muitos algoritmos para problemas em grafos podem utilizar o seguinte esquema recursivo: dado um grafo conexo G , particionar G em dois subgrafos conexos de tamanhos similares, resolver os subproblemas correspondentes para esses subgrafos e de suas soluções obter uma solução para o grafo original G . Neste

contexto, é desejável que a bipartição seja tão balanceada quanto possível, minimizando assim a profundidade da recursão. Além disso, no contexto de processamento distribuído, do ponto de vista de balanceamento de carga, partições balanceadas são relevantes.

Problemas de partições balanceadas podem ser mais gerais, no sentido que o grafo de entrada pode ter pesos associados aos seus vértices, e o balanceamento deve levar em conta o peso dos subgrafos conexos da partição. Ao precisar esse conceito de balanceamento, obtemos algumas variantes deste problema. A seguir, focamos em um problema dessa natureza, que chamamos de Problema da *Partição Conexa Balanceada (PCB)*, definido como segue.

Dado um inteiro $k \geq 2$, um grafo conexo $G=(V,A)$ com peso não-negativo $w(v)$ definido para cada vértice $v \in V$, encontrar uma k -partição (V_1, V_2, \dots, V_k) de V tal que $G[V_i]$ (subgrafo induzido por V_i) seja conexo para $1 \leq i \leq k$, e o peso do subgrafo mais leve seja o maior possível. Mais formalmente, queremos encontrar uma tal partição que maximiza a função $\min\{w(V_i): i=1, \dots, k\}$, onde $w(X)$ denota o peso do conjunto X , definido como a soma dos pesos dos elementos em X .

A variante do PCB quando k é fixo (não faz parte da entrada) é denotada por PCB k . Em 1983, Camerini, Galbiati e Maffioli [1] mostraram que PCB2 é NP-difícil, donde se obtém que PCB e PCB k para $k \geq 3$ também são NP-difíceis. Dentre as aplicações do PCB k , mencionamos problemas em processamento de imagens, clustering,

logística, sistemas operacionais e robótica.

Nosso interesse pelos problemas PCB e PCB k surgiu no início de 2000, com orientações de um doutorado e um mestrado. Os primeiros estudos tiveram como foco o desenvolvimento de algoritmos de aproximação e questões mais teóricas sobre limites de aproximabilidade. Dizemos que um algoritmo para um problema de maximização é uma *r*-aproximação se, para qualquer instância desse problema, encontra em tempo polinomial uma solução viável cujo valor é pelo menos *r* vezes o valor de uma solução ótima para essa instância. Na década de 90 já era conhecido um algoritmo de 3/4-aproximação, obtido por Chlebíkova [3] para o PCB 2 , e um algoritmo polinomial exato para o PCB k devido a Perl e Schach no caso especial em que o grafo de entrada é uma árvore. A tese de doutorado de Salgado (2004) apresenta alguns resultados de inaproximabilidade e algoritmos de aproximação para grafos *k*-conexos, $k \in \{3, 4\}$, e também resultados para o caso em que as classes devem ter uma certa proporção fixa do peso total [6]. Em colaboração com Chataigner e Salgado [2] mostramos que o PCB (em que *k* faz parte da entrada) não pode ser aproximado por uma razão melhor que 5/6 sob a hipótese que $P \neq NP$. Em sua dissertação de mestrado, Lucindo [4] implementou (e disponibilizou para acesso público) algoritmos com bom desempenho empírico baseados na geração aleatória de árvores geradoras, e uso de algoritmos polinomiais para árvores.

Mais recentemente, investigamos o PCB k sob o ponto de vista poliédrico. Nesta abordagem, o conjunto das soluções viáveis são pontos no espaço de dimensão

[*V*], e o que se busca é a descrição linear -- ainda que parcial -- do poliedro definido como o fecho convexo desses pontos. Essa abordagem foi feita com vistas à implementação de algoritmos exatos de *branch-and-bound* e *branch-and-cut*. Esses estudos, conduzidos em colaboração com Miyazawa, Ota e Moura [5], propõem formulações para o PCB k tanto como Programa Linear Misto (PLM) e como Programa Linear Inteiro (PLI).

Consideramos duas formulações como PLM, ambas baseadas em fluxos em redes, e com um número polinomial de variáveis e restrições. A formulação como PLI tem somente variáveis binárias e um número potencialmente grande de desigualdades. Neste último modelo, usamos algoritmos de separação polinomiais para resolver o PL relaxado. Também fortalecemos o modelo com novas desigualdades válidas, para as quais projetamos algoritmos de separação polinomiais.

Essas formulações impõem uma ordenação das classes da partição, de acordo com seus pesos, e consideram como objetivo maximizar o peso da classe mais leve. Tais formulações permitem também considerar outras funções-objetivo que capturam o conceito de balanceamento como, por exemplo, minimizar o peso de uma classe mais pesada ou minimizar a maior diferença de peso entre as classes.

Os experimentos computacionais conduzidos mostraram que os algoritmos exatos que implementamos têm desempenhos substancialmente melhores que os métodos exatos publicados na literatura. Além dos dados da literatura, consideramos também dados obtidos em órgãos públicos sobre a quantidade de

delitos em regiões urbanas (como Chicago e Campinas) para definir instâncias do PCB k úteis para logísticas de patrulhamento policial dessas regiões.

Adicionalmente, obtivemos os primeiros resultados poliédricos associados à segunda formulação restrita ao caso de pesos uniformes. Tais resultados podem levar a novas abordagens para o PCB k . Devido ao espaço limitado, deixamos de

mencionar esses resultados, referências a outros trabalhos, e análises comparativas do desempenho dos algoritmos que propusemos e dos existentes na literatura (todos testados no mesmo ambiente computacional). Parte desses resultados podem ser vistos em [5].

Referências

1. CAMERINI, P; GALBIATI, G.; MAFFIOLI, F. On the complexity of finding multi constrained spanning trees. *Discrete Applied Mathematics*, v. 5, p. 39-50, 1983.
2. CHATAIGNER, F.; SALGADO, L. R. B.; WAKABAYASHI, Y. Approximation and inapproximability results on balanced connected partitions of graphs. *Discrete Mathematics & Theoretical Computer Science*, 39 (1), 2007 (eletrônico).
3. CHLEBÍKOVA, J. Approximating the maximally connected partition problem in graphs. *Information Processing Letters*, v. 60, p. 225-230, 1996.
4. LUCINDO, R. P. F. L. Partições de grafos em subgrafos conexos balanceados. Dissertação de mestrado, IME-USP, 2007. <https://www.ime.usp.br/~lucindo/graphpar/>
5. MIYAZAWA, F. K.; MOURA, P. F. S.; OTA, M. J.; WAKABAYASHI, Y. Cut and flow formulations for the balanced connected k-partition problem. Em: Baiou, M.; Gendron, B.; Günlük, O.; Mahjoub, A. (editores). *Combinatorial Optimization. ISCO 2020. Lecture Notes in Computer Science*, v. 12176, p. 128-139, Springer.
6. SALGADO, L. R. B.; WAKABAYASHI, Y. Approximation results on balanced connected partitions of graphs. *Proceedings of LACGA, 2004. Electronic Notes in Discrete Mathematics* 18 (2004) p. 207-212.



YOSHIKO WAKABAYASHI é Professora Titular do Departamento de Ciência da Computação da Universidade de São Paulo. Obteve o doutorado em Matemática Aplicada na Universität Augsburg, Alemanha, em 1986. Faz pesquisa nas áreas de otimização combinatória e teoria dos grafos. Em 2010 foi admitida na Ordem Nacional do Mérito Científico, classe Comendador. É membro da Academia de Ciências do Estado de São Paulo desde 2012, e membro da Academia Brasileira de Ciências, na área de Matemática, desde 2019.

ASSOCIE-SE À SBC

E NOS AJUDE A ESTIMULAR O ACESSO À INFORMAÇÃO
E À CULTURA POR MEIO DA COMPUTAÇÃO.

CATEGORIA	INVESTIMENTO
EFETIVO/FUNDADOR	R\$ 260,00
EFETIVO ASSOCIADO À ACM	R\$ 234,00
EFETIVO PROFESSOR DE EDUCAÇÃO BÁSICA	R\$ 91,00
ESTUDANTE PÓS-GRADUAÇÃO	R\$ 91,00
ESTUDANTE PÓS-GRADUAÇÃO ASSOCIADO À ACM	R\$ 87,00
ESTUDANTE	R\$ 22,00
INSTITUCIONAL*	R\$ 2.350,00

ASSOCIADOS SBIS E SBIAGRO: 10% DE DESCONTO, MEDIANTE COMPROVAÇÃO.

ASSOCIADOS NA CATEGORIA EFETIVO ABE, SBEM, SBHMAT, SBM, SBMAC: 50% DE DESCONTO, MEDIANTE COMPROVAÇÃO.
SOLICITE ENVIANDO E-MAIL PARA SBC@SBC.ORG.BR.

*DEVIDO À PANDEMIA, A SBC ESTÁ APLICANDO O MESMO VALOR DE 2019 PARA A CATEGORIA INSTITUCIONAL.

PARA MAIS INFORMAÇÕES ACESSE:
SBC.ORG.BR