

BOAS PRÁTICAS



Inteligência artificial na revisão por pares

Software aponta possíveis conflitos de interesse envolvendo autores de artigos e pesquisadores que avaliam sua qualidade

A editora suíça Frontiers, que publica mais de 90 periódicos científicos em acesso aberto, desenvolveu um software que utiliza inteligência artificial para examinar artigos e identificar até 20 diferentes problemas relacionados à sua integridade, como plágio ou imagens com sinais de manipulação. O programa chamou a atenção, no entanto, por oferecer um serviço inovador: ele avisa quando os autores de um manuscrito e os editores e revisores que estão apreciando o seu conteúdo já assinaram juntos outros artigos, em uma proximidade que pode tornar a avaliação subjetiva e configurar um conflito de interesses. Denominado Artificial Intelligence Review Assistant (Aira), a ferramenta está sendo utilizada no processo de revisão por pares de revistas da Frontiers com o propósito de fornecer um alerta objetivo sobre interações prévias entre quem produz e quem avalia o conhecimento. Cabe a um juiz de carne e osso – o editor-chefe do periódico, em última instância – arbitrar se esses contatos anteriores afetam ou não a isenção da análise. “É necessário ter apoio tecnológico para verificar em larga escala conflitos de interesse entre autores e revisores”, disse o cientista da computação Daniel Petriariu, diretor de Desenvolvimento de Produtos da Frontiers, no lançamento do programa, em julho.

O Committee on Publication Ethics (Cope), fórum que discute temas relacionados à integridade na ciência, define conflito de interesses como uma situação em que pesquisadores ou instituições científicas têm interesses que concorrem entre si, de caráter profissional ou econômico, de modo que tomar partido de um deles pode comprometer a imparcialidade de suas decisões. A existência de um conflito de interesses, observa o Cope, não é determinante para a ocorrência de má conduta: a despeito de tais circunstâncias, um estudo pode perfeitamente ter resultados válidos e fidedignos. Mas, quando isso acontece, é obrigatório lidar com o problema de forma transparente. Por isso, exige-se que pesquisadores assinem declarações de conflitos de interesse sempre que apresentar um projeto ou publicar um artigo. Fica a cargo do avaliador do projeto ou do leitor do *paper* ponderar os resultados tendo em mente os interesses envolvidos.

Um caso típico de má conduta é esconder que recebeu financiamento de uma empresa para realizar uma pesquisa em cujos resultados ela tem interesse econômico. Foi o caso do oncologista José Baselga, que perdeu o cargo de diretor clínico do Memorial Sloan Kettering Cancer Center, em Nova York, depois de reconhecer que omitiu ligações com empresas farmacêuticas em dezenas de artigos científicos de sua autoria (*ver* Pesquisa FAPESP nº 272). Também há situações de conflitos de interesse que macularam a objetividade da revisão por pares. Em 2017, a revista *Scientific World Journal*, da editora Hindawi, anunciou a reatuação de dois artigos quando se soube que o autor dos trabalhos, Zheng Xu, era um colaborador frequente de Xiangfeng Luo, editor do periódico e pesquisador da Universidade de Shanghai. Eles assinaram juntos dezenas de artigos e Luo foi orientador de doutorado de Xu.

O objetivo do software Aira é detectar uma fonte potencial de conflito de interesses que poderia passar despercebida. Mas seu alcance é limitado – ele só aponta problemas quando há dados disponíveis para corroborá-los. E parte desses conflitos ocorre em uma zona de sombras. Se o autor e o avaliador receberem financiamento de um mesmo patrocinador e omitirem esse vínculo, não é o software que irá revelar isso.

Um documento sobre o futuro da revisão por pares publicado em 2017 pela editora BioMed Central mostrou que a inteligência artificial vem sendo utilizada de várias formas na avaliação de artigos científicos. Há editoras que usam tais recursos para rastrear a produção científica disponível on-line a fim de identificar pesquisadores com perfil para atuar como revisores de papers em suas disciplinas. Da mesma forma, a inteligência artificial está municando novos softwares capazes de detectar plágio. Eles identificam quando frases ou parágrafos foram reescritos para driblar algoritmos que só conseguem captar cópias literais.

O exemplo mais polêmico foi o do software Statcheck, criado em 2015 por pesquisadores da Universidade de Tilburg, nos Países Baixos (*ver* Pesquisa FAPESP nº 253), e hoje incorporado ao dia a dia de centenas de revistas. Uma controvérsia envolveu não a tecnologia, capaz de refazer os cálculos descritos em

artigos científicos já publicados e apontar erros estatísticos, mas a forma como ela foi utilizada. Cinquenta mil artigos da área de psicologia foram analisados pelo software e os resultados desse escrutínio foram divulgados na plataforma PubPeer, expondo em público pequenos e grandes erros cometidos por milhares de autores. Mais de 25 mil artigos apresentaram alguma inconsistência.

Para a psicóloga Michèle Nuijten, professora da Universidade de Tilburg e uma das criadoras do Statcheck, ferramentas como o Aira podem ser valiosas na avaliação por pares. “Precisamos buscar soluções inovadoras que ajudem a aumentar a qualidade e a robustez dos estudos científicos, e a inteligência artificial tem um papel a desempenhar nisso”, disse, em reportagem sobre o Aira publicada no jornal *The New York Times*. Mas Nuijten adverte que essas ferramentas devem ser usadas apenas no apoio ao trabalho dos editores e que cabe ao homem, e não à máquina, decidir se um *paper* deve ou não ser publicado. “Fico preocupada com a ideia de que um artigo possa ser rejeitado apenas porque uma ferramenta de inteligência artificial identificou um problema, sem que se verifique realmente o que está acontecendo.”

Agências de fomento também têm normas para prevenir conflitos de interesse na avaliação de projetos. A FAPESP há muito tempo solicita que assessores escolhidos para analisar propostas declarem se têm interesses relacionados a elas. Uma verificação prévia dos nomes de potenciais assessores já descarta os que têm vínculos como parentesco, afiliação à mesma instituição e participação no projeto. Em maio, começou a funcionar um sistema automático de identificação de possíveis revisores desenvolvido pela Gerência de Informática da Fundação, que sugere aos coordenadores de área uma lista de 20 pesquisadores capazes de avaliar cada projeto. O sistema de busca compara parâmetros como títulos e palavras-chave dos projetos com uma base de perfis de assessores e aponta, por ordem de relevância, as melhores combinações. A tecnologia responsável pela classificação das palavras utilizou recursos de inteligência artificial. Além de tornar a seleção mais objetiva, o sistema checa de modo automático se os revisores têm vínculos com os proponentes. ■ **Fabício Marques**

Paper de 1951 sobre homossexualidade sofre retratação

O *Journal of Nervous and Mental Disease*, periódico de neuropsiquiatria que circula desde 1874, anunciou a retratação de um artigo científico de quase 70 anos de idade. Apresentado em 1948 em um evento de uma associação de psiquiatras e publicado na revista em 1951, o trabalho descreve as observações sobre o comportamento de estudantes homossexuais que receberam atendimento psicológico na Universidade de Wisconsin-Madison, nos Estados Unidos, feitas por Benjamin Glover, professor de psicologia da instituição. No *paper*, Glover descreve os alunos como “uma paródia e um paradoxo emocional” e diz que eles são “uma tragédia contínua no fracasso para encontrar gratificação sexual”. Ainda de acordo com o autor, os estudantes gays seriam “devotados aos seus amores com uma paixão expressa; no entanto, eles têm pouco ou nenhum sentimento pelos pais e duvidou que ficariam chateados se a morte ou uma doença grave os atingisse”. Outra amostra: “Há um egoísmo narcisista em

seu desprezo pelas pessoas como um todo, nenhum sentimento nacionalista ou patriótico, um desprezo geral pelos valores sociais da lei, da religião ou o aperfeiçoamento da humanidade”.

O editor da revista, John Talbot, definiu as opiniões de Glover como “crenças que consideramos abomináveis atualmente”, mas lembrou que não se tratava de um caso isolado e que os periódicos da época publicaram artigos defendendo conceitos igualmente inaceitáveis, como a eugenia. Por isso, determinou que o artigo de Glover, mesmo com o carimbo de retratação, continue disponível como um testemunho de seu tempo. Antecipando-se às críticas de que estaria julgando o pesquisador sem considerar o conhecimento disponível da época, a revista publicou também um editorial do neurocientista Simon LeVay, famoso por estudos sobre as origens biológicas e comportamentais da homossexualidade. Foi LeVay quem pediu a retratação do artigo.

No texto, o pesquisador explicou que o espírito do tempo não é suficiente pa-

ra explicar a leitura preconceituosa de Benjamin Glover. LeVay reconhece que autores da época, como Edmund Bergler, que lançou em 1956 o livro *Homossexualidade: Doença ou modo de vida?*, partilhavam as mesmas crenças. Mas observa que, em contraste, vários outros pesquisadores adotaram uma abordagem mais científica e objetiva. “Em 1958, por exemplo, o psiquiatra Mathew Ross e o estudante de medicina Fred Mendelsohn publicaram um estudo com 133 estudantes homossexuais que fizeram psicoterapia no serviço de saúde estudantil da Universidade da Califórnia, em Los Angeles, juntamente a um grupo de controle de estudantes heterossexuais. Ao contrário de Glover, esses autores não expressaram opiniões negativas sobre seus alunos homossexuais, fizeram pouco esforço para mudar a orientação sexual deles e descobriram que estudantes homossexuais e heterossexuais eram semelhantes em muitos aspectos e igualmente receptivos à psicoterapia”, escreveu LeVay.

Ambições frustradas

Um grupo de pesquisadores dos Países Baixos dedicou-se, nos últimos meses, a preparar um grande levantamento multidisciplinar sobre integridade científica em universidades e instituições de pesquisa nacionais: mais de 40 mil pesquisadores foram convidados a preencher um questionário on-line sobre tópicos como práticas questionáveis e má conduta. Os resultados, contudo, ficaram muito aquém das ambições iniciais, pois apenas 15% aceitaram participar e enviaram respostas até a data-limite, 7 de dezembro.

De 15 universidades convidadas, apenas cinco colaboraram com o estudo. Lex Bouter, pesquisador da Universidade Livre de Amsterdã e idealizador da iniciativa, assegurou aos dirigentes de universidades que o levantamento não seria usado para abastecer nenhum ranking de mau comportamento, mas eles se queixaram da ênfase excessiva em te-

mas relacionados à má conduta. “Achei que seria tendencioso”, disse à revista *Science* Henk Kummeling, presidente da Universidade de Utrecht, uma das que se recusaram a participar. “Se você perguntar apenas sobre práticas de pesquisa questionáveis, já sabe de antemão o que vai ganhar como resposta.”

O rol de perguntas foi ampliado para tentar deixar o estudo mais palatável, incluindo assuntos como compartilhamento de dados e ciência aberta, mas o esforço não conseguiu derrubar as resistências. Jeroen de Ridder, filósofo da ciência da Universidade Livre de Amsterdã sem envolvimento com o estudo, se disse desapontado com o boicote e a perda de oportunidade de analisar os problemas relacionados à integridade científica no país. Ele nega que a pesquisa tenha falhas metodológicas: “É a mais cuidadosa e completa que se poderia desejar”, afirmou.