

# Promessa de



Cientistas da computação propõem veto a nova revista de inteligência artificial por ela não adotar acesso aberto

Cerca de 3 mil pesquisadores e estudantes de computação assinaram em maio um manifesto se comprometendo a não submeter artigos a um periódico que deve ser lançado no início de 2019 nem participar de seu processo de revisão por pares. A revista em questão é a *Nature Machine Intelligence*, do grupo Springer Nature, que pretende divulgar estudos sobre inteligência artificial, aprendizado de máquina e robótica. A promessa de boicote tem uma justificativa simples: a produção científica em ciência da computação, e particularmente em inteligência artificial, segue um modelo peculiar de publicação em que as descobertas são divulgadas e debatidas em conferências

e, em seguida, arquivadas em repositórios de acesso aberto – o ArXiv é uma destinação frequente dos manuscritos. Esse paradigma se contrapõe ao sistema tradicional de comunicação científica, em que o acesso é disponível para quem se dispuser a pagar para ler, e também a um padrão mais moderno que ganhou espaço nos últimos tempos, no qual o acesso é aberto, mas o autor precisa pagar uma taxa para financiar a publicação.

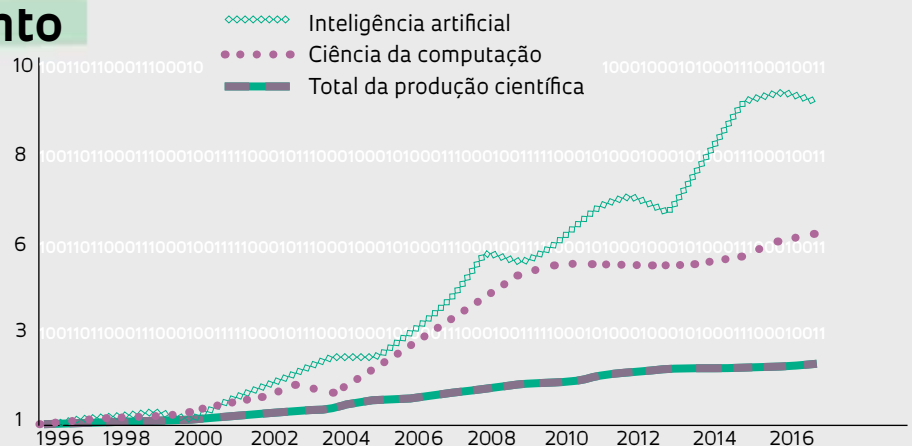
“Não vemos papel algum, no futuro da pesquisa em aprendizado de máquinas, para revistas de acesso fechado ou que cobram taxas de autores”, informa o manifesto. “Em contraste, receberíamos de bom grado novas conferências e revistas com publicação a custo ze-

ro.” A declaração foi subscrita tanto por pesquisadores ligados a instituições de pesquisa, entre as quais a Universidade Harvard e o Massachusetts Institute of Technology (MIT), nos Estados Unidos, quanto por especialistas vinculados a empresas como Google, Microsoft, Facebook e IBM. A ausência de nomes da Apple foi sentida.

Entre os signatários, destacam-se nomes como os dos cientistas da computação Yann LeCun, professor da Universidade de Nova York e diretor do braço de pesquisa em inteligência artificial do Facebook, e Yoshua Bengio, da Universidade de Montreal, no Canadá, pioneiros na pesquisa sobre redes neurais. “A comunidade de aprendizado de máquina

## Ritmos de crescimento

Quantas vezes cresceu, entre 1996 e 2016, o total da produção científica e o número de artigos publicados sobre inteligência artificial e ciência da computação. Publicações em 1996 equivalem a 1



FONTE: REVISTA SCIENCE E SCOPUS

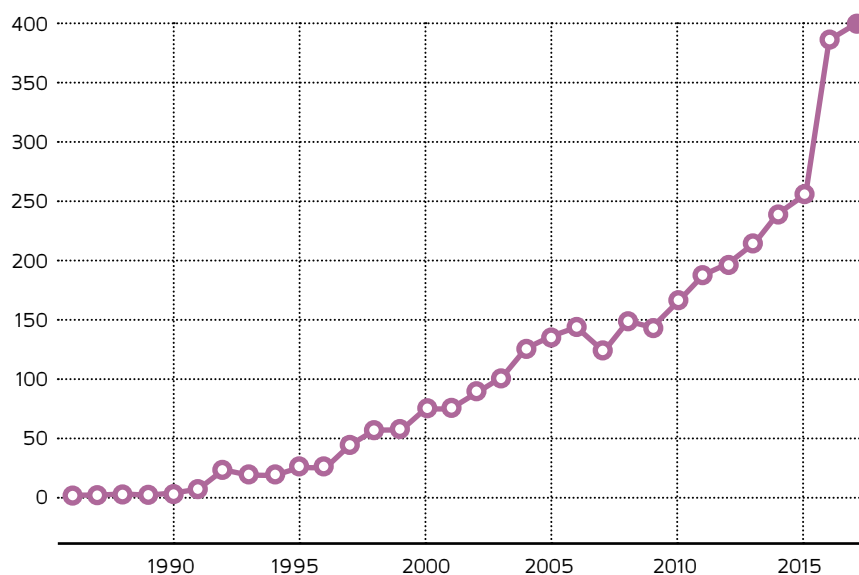
tem feito um grande trabalho ao publicar pesquisas em acesso aberto; reverter essa tendência utilizando um modelo fechado não parece ser uma boa ideia”, disse, segundo a revista *Forbes*, Jeff Dean, diretor do Google AI, braço de pesquisa do Google sobre inteligência artificial, outro que assinou a declaração. Segundo dados da base Scopus compilados pela revista *Science*, o número de trabalhos científicos em inteligência artificial multiplicou-se por 10 entre 1996 e 2016, acima da média da ciência da computação, que cresceu seis vezes no mesmo período. O Brasil segue a mesma tendência: há pesquisadores publicando na área desde os anos 1990. Eles formaram uma força de pesquisa responsável por uma produção que mais que dobrou de tamanho entre 2010 e 2016 (ver gráfico nesta página).

**O** responsável pelo manifesto, o cientista da computação Thomas Dietterich, que é professor emérito da Universidade do Estado de Oregon, nos Estados Unidos, usou sua conta no Twitter para convidar colegas a cancelar a declaração. Recebeu em resposta um tweet da própria revista *Nature Machine Intelligence*, em um tom conciliador: “Consideramos que a *Nature Machine Intelligence* pode coexistir e fornecer um serviço – para os que estiverem interessados –, conectando diferentes campos, oferecendo um canal para trabalhos interdisciplinares e conduzindo um processo rigoroso de avaliação por pares”. A porta-voz da *Springer Nature* em Londres, Susie Winter, divulgou um comunicado defendendo o modelo: “Acreditamos que a maneira mais justa de produzir revistas altamente seletivas como essa e garantir sua sustentabilidade no longo prazo e para uma comunidade a mais ampla possível é distribuir os custos entre muitos leitores”.

A repercussão do manifesto foi grande, por mobilizar um campo emergente de pesquisa na defesa do acesso aberto e também por contrapor-se a um dos maiores grupos de comunicação científica do mundo. Mas não chegou a surpreender quem acompanha a evolução do conhecimento em inteligência artificial e o ambiente em que ele é produzido. Um round anterior desse embate foi travado em 2001, quando a cientista da computação Leslie Kaelbling, do MIT, lançou o periódico de acesso aberto *Journal of*

## Avanço do conhecimento científico

Evolução do número de publicações de autores do Brasil sobre tópicos ligados à pesquisa em inteligência artificial\*



\* Inteligência artificial, aprendizado de máquina, aprendizagem profunda e redes neurais

FONTE: WEB OF SCIENCE

## Dados integrados

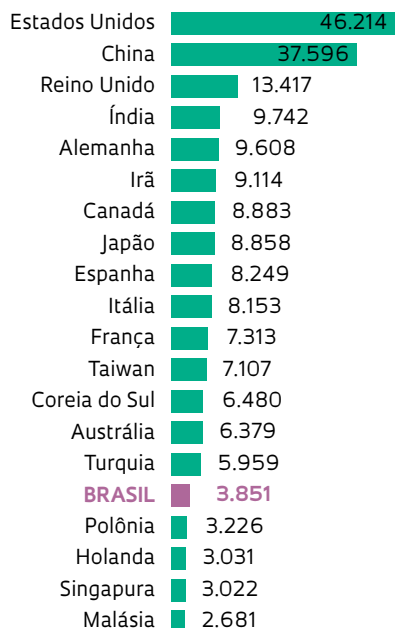
Sistema reúne informações sobre a produção de pesquisadores brasileiros em subáreas da ciência da computação

Pesquisadores das universidades federais de Minas Gerais (UFMG) e de Uberlândia (UFU) criaram um sistema que fornece dados comparativos sobre a produção dos departamentos de ciência da computação de instituições de ensino superior do Brasil. Lançado em janeiro, o CSIndex (<http://goo.gl/qWuS2w>) oferece informações que podem ser úteis tanto para avaliar o desempenho de grupos de pesquisa quanto para orientar estudantes na escolha de cursos de mestrado e doutorado. O serviço reúne e organiza a produção total dos departamentos em 16 subáreas, gerando um escore para cada um. O Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP) detém o escore mais alto na

área de inteligência artificial. Já em engenharia de software, a Universidade Federal de Pernambuco (UFPE) conta com a melhor nota, enquanto o Departamento de Ciência da Computação da UFMG tem o escore mais elevado na área de banco de dados.

A coleta das informações é feita por meio do acompanhamento de artigos apresentados em conferências indexadas em bases internacionais. Na ciência da computação, a publicação de trabalhos em conferências costuma ser mais importante do que a publicação de artigos em revistas científicas. O CSIndex monitora artigos publicados nos anais de 162 conferências, distribuídas em subáreas como engenharia de software,

## Número de publicações por país na base Web of Science sobre tópicos ligados à pesquisa em inteligência artificial\*



*Machine Learning Research (JMLR)* como alternativa à prestigiosa revista *Machine Learning*, do Kluwer Academic Press, de acesso restrito. À época, boa parte do conselho editorial da *Machine Learning* renunciou e se transferiu para a *JMLR*.

A percepção de que o processo tradicional de publicação científica era demorado para um campo do conhecimento em efervescência levou os cientistas da computação a adotar um modelo em que conferências e repositórios de acesso aberto têm peso preponderante e a revisão por pares é aberta à contribuição de qualquer pesquisador e feita após a divulgação dos resultados. Esse modelo continua em evolução. Em 2013, o cientista da computação Andrew McCallum, da Universidade de Massachusetts, em Amherst, nos Estados Unidos, criou o site Open Review, por meio do qual autores podem publicar manuscritos apresentados em conferências e convidar colegas para comentá-los. Em pouco tempo, as principais conferências sobre inteligência artificial passaram a utilizar os serviços do site.

Rui Seabra Ferreira Júnior, presidente da Associação Brasileira de Editores Científicos, observa que os modelos de publicação científica passam por uma fase de transição e ainda não há clareza sobre o que deverão se tornar no futu-

ro. “A revisão por pares existe há mais de 300 anos e não há evidência de que vá perder importância. Mas há alguns campos disciplinares mais teóricos, como algumas áreas da física e da ciência da computação, em que é possível um pesquisador apresentar conceitos novos em um repositório de acesso aberto e abri-los para discussão entre os colegas, cabendo a eles validá-los ou não”, afirma. “Já em áreas aplicadas, a revisão por pares é essencial. Veja o caso da biologia ou da medicina: seria necessário que os colegas reproduzissem os experimentos descritos em repositórios para poder validá-los, o que seria inviável.”

O físico Paul Ginsparg, pesquisador da Universidade Cornell e fundador do repositório ArXiv, elogiou o manifesto dos cientistas da computação, mas se mostrou cético em relação à sua capacidade de influenciar o modelo de publicação científica como um todo. “Pessoalmente, não tenho nenhuma animosidade contra o modelo de assinatura”, afirmou à revista *Science*. Ele considera que os signatários do manifesto têm uma visão pouco realista sobre a possibilidade de publicar a custo zero. “Servidores de repositórios são baratos, mas um controle sistemático de qualidade exige mão de obra intensiva e custa dinheiro.” ■ **Fabício Marques**

linguagens de programação, arquitetura de computadores e bancos de dados.

O sistema usa um algoritmo projetado para recuperar informações da Digital Bibliography & Library Project (DBLP), um dos principais repositórios bibliográficos de ciência da computação, hospedado na Universidade de Trier, na Alemanha. O DBLP indexa mais de 3 milhões de artigos de 1,7 milhão de autores e é atualizado mensalmente. “O CSIndex filtra os trabalhos com autores brasileiros, classificando-os segundo as subáreas”, explica o cientista da computação Marco Tulio Valente, da UFMG, um dos criadores do sistema, ao lado do também cientista da computação Klérison Paixão, da UFU. Em seguida, calcula o escore para cada departamento. O CSIndex monitora a produção científica de 818 cientistas da computação de universidades públicas e privadas, e de institutos federais do país.

As conferências de maior prestígio de ciência da computação têm um processo de seleção bastante rigoroso e em geral aceitam menos de 20% dos artigos submetidos. Os trabalhos publicados em 34 dessas conferências mais disputadas receberam peso 1, enquanto os divulgados nas demais têm peso 0,33. “O escore de cada departamento é o resultado da combinação do número de artigos publicados pelos pesquisadores em conferências ‘top’ e nas demais conferências monitoradas pelo CSIndex”, explica Valente.

A ideia, segundo ele, é que o CSIndex constitua uma plataforma independente, podendo complementar os sistemas correntes de avaliação, como o realizado a cada quatro anos pela Coordenação Nacional de Aperfeiçoamento de Pessoal de Nível Superior (Capes), que monitora a qualidade dos programas de pós-graduação. Na concepção do

cientista da computação André Sampaio Gradvohl, da Faculdade de Tecnologia da Universidade Estadual de Campinas (Unicamp), é importante que o CSIndex considere outros repositórios além do DBLP. “A ampliação do número de repositórios é recomendável para considerar trabalhos publicados em periódicos científicos”, diz o pesquisador, que não participou da concepção do sistema.

Para o cientista da computação Filipe Saraiva, da Universidade Federal do Pará, também não envolvido na criação do CSIndex, a chance de acompanhar a qualidade da produção dos outros departamentos abre portas para novas colaborações. “É interessante observar métricas relacionadas às pesquisas em ciência da computação desenvolvidas no Brasil, ainda mais quando se baseiam em um filtro que prioriza publicações de alto impacto”, avalia. ■ **Rodrigo de Oliveira Andrade**