

A arte de fazer parceiros

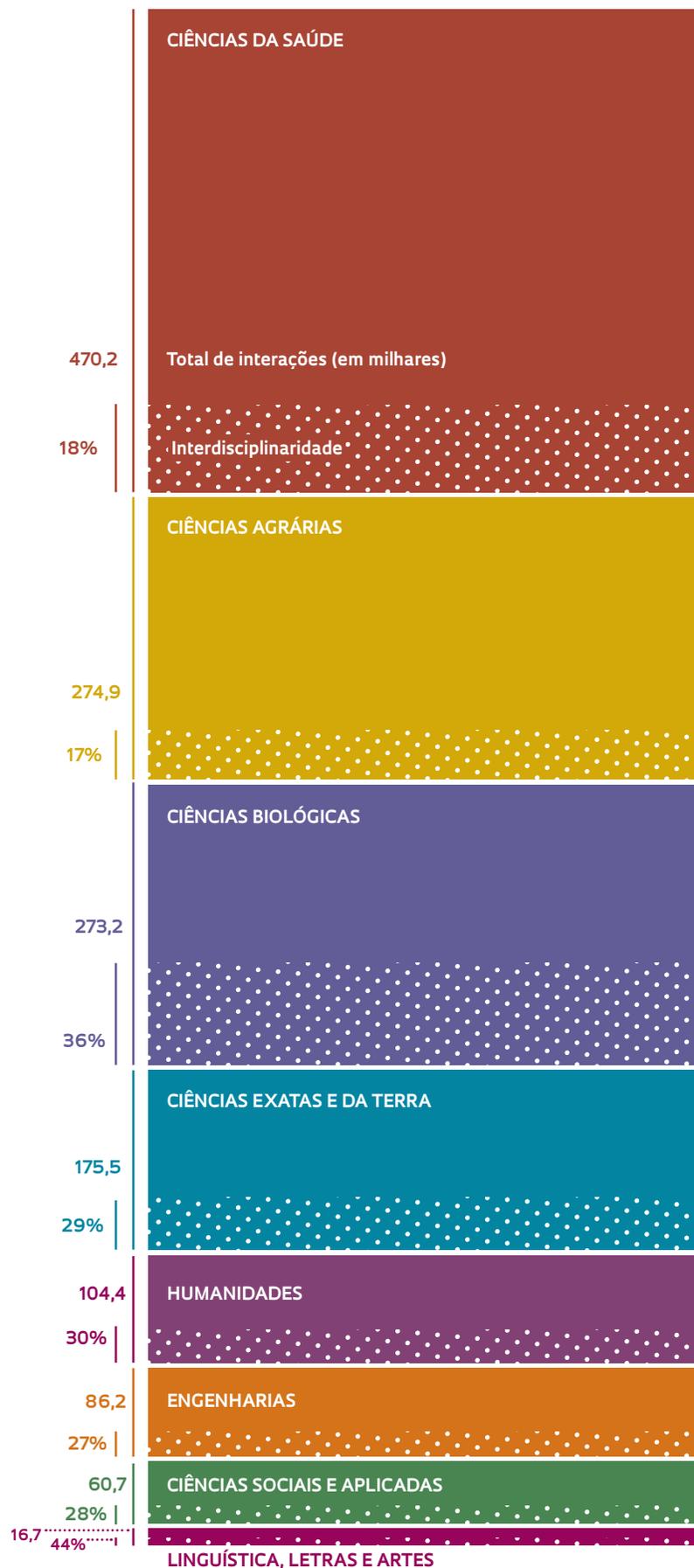
Cruzamento de dados de 1,1 milhão de pesquisadores brasileiros mostra que colaborações multiplicaram-se nos últimos 20 anos

Bruno de Pierro

Pesquisadores das universidades de São Paulo (USP), Federal do ABC (UFABC) e Tecnológica Federal do Paraná debruçaram-se sobre uma massa gigantesca de informações – a produção bibliográfica de cerca de 1,1 milhão de pesquisadores e estudantes brasileiros que estão registrados na plataforma de currículos Lattes – e obtiveram um panorama inédito das colaborações celebradas entre eles, observadas por meio de coautoria de artigos científicos, livros e capítulos de livros. O trabalho, publicado em janeiro no *Journal of the Association for Information Science and Technology*, mostra que o número de colaborações entre pesquisadores brasileiros teve um avanço notável nas últimas duas décadas. Entre 2008 e 2010 foi registrado pouco mais de 1 milhão de colaborações, em contraste com as 63.944 observadas entre 1990 e 1992. Esse avanço supera largamente o crescimento da comunidade científica no período.

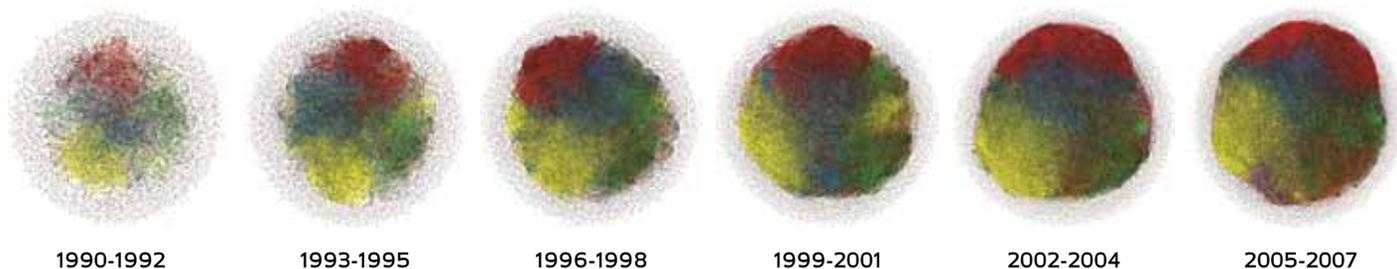
Os autores da pesquisa analisaram a evolução de oito grandes áreas do conhecimento (*ver gráfico*). Ciências da saúde foi um dos destaques: quase 42% dos 272.783 pesquisadores estabele-

TOTAL DE INTERAÇÕES ENTRE PESQUISADORES BRASILEIROS POR GRANDE ÁREA (1990-2010), EM MILHARES, E A PORCENTAGEM DE PARCERIAS INTERDISCIPLINARES



O adensamento das colaborações

As imagens são a representação gráfica das redes de colaboração de pesquisadores brasileiros entre 1990 e 2010



- Ciências agrárias
- Ciências biológicas
- Ciências exatas e da Terra
- Humanidades
- Ciências sociais aplicadas
- Ciências da saúde
- Engenharias
- Linguística, letras e artes
- Mais de uma grande área

ceram colaborações com colegas da mesma área. Na outra ponta está linguística, letras e artes: apenas 15% dos quase 100 mil currículos avaliados exibiam coautorias entre pesquisadores da área. Apesar disso, linguística, letras e artes têm uma peculiaridade: seus pesquisadores foram os que estabeleceram mais parcerias com colegas de outras áreas, especialmente a de humanidades. “Em termos relativos, linguística, letras e artes é a área com maior grau de interdisciplinaridade. Já em números absolutos, ciências biológicas contabilizou o maior número de interações com outras áreas”, diz Jesús Pascual Mena-Chalco, professor do Centro de Matemática, Computação e Cognição da UFABC e autor principal da pesquisa.

Os dados referentes às colaborações nas áreas de ciências humanas surpreenderam Samile Vanz, professora da Faculdade de Biblioteconomia da Universidade Federal do Rio Grande do Sul. As colaborações entre 1990 e 2010 alcançaram 169.869 na área de humanidades e 102.058 na de ciências sociais aplicadas. Em outros trabalhos, ela observa, a produção científica dessas áreas parece mais tímida. “Isso acontece porque os estudos geralmente utilizam as bases de dados Web of Science e Scopus, que indexam principalmente artigos científicos. Os pesquisadores das áreas de humanidades tradicionalmente publi-

cam bastante em livros e capítulos de livros e isso é mais visível nos dados do currículo Lattes”, diz Samile.

A contribuição do estudo não se restringe a caracterizar a colaboração entre cientistas brasileiros. Ele também produziu avanços no campo da computação e da tecnologia da informação, ao desenvolver modelos matemáticos e algoritmos capazes de refinar a organização de grandes quantidades de dados. “O que fizemos foi olhar para o universo que é a Plataforma Lattes e, por meio de novos algoritmos, identificar as colaborações entre os autores”, diz Mena-Chalco. “A partir dessa ferramenta metodológica, sociólogos e gestores poderão fazer análises mais profundas”, sugere. “É um trabalho pioneiro”, diz Rogério Meneghini, coordenador científico da biblioteca virtual SciELO Brasil. “Os autores tiveram a capacidade de extrair dados da Plataforma Lattes e, para isso, desenvolveram uma metodologia própria. Ainda é difícil extrair informações dessa plataforma.” Atualmente, a plataforma, organizada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), conta com mais de 2,7 milhões de currículos e, por reunir informações de praticamente todos os pesquisadores brasileiros, é considerada uma base de dados única no mundo.

BIG DATA

Para Roberto Marcondes Cesar Júnior, professor do Instituto de Matemática e Estatística da USP e outro autor do estudo, a pesquisa relaciona-se diretamente com um tema emergente, o Big Data – um conceito da tecnologia de informação baseado no armazenamento e na interpretação de gigantescos volumes

de dados. Nos últimos anos, o grupo de Marcondes, que é coordenador-adjunto de Ciências Exatas e Engenharias da FAPESP, tem se dedicado ao tema com intensidade. Esse estudo faz parte de uma linha de pesquisa mais ambiciosa dentro do Programa de Apoio a Núcleos de Excelência (Pronex), fruto de um convênio entre a FAPESP e o CNPq. “Um dos problemas pensados por nós é como criar o modelo matemático que permita fazer perguntas do tipo: se eu contratar um pesquisador hoje, qual é a contribuição esperada de sua produção científica para o meu departamento no futuro? O advento do Big Data vem da percepção de que temos mais dados para analisar do que somos capazes de dar conta humanamente”, diz Marcondes.

Para Jacqueline Leta, professora da Universidade Federal do Rio de Janeiro, os resultados da pesquisa ajudam a confirmar algumas hipóteses. “A indicação de que ciências biológicas e ciências da saúde são as áreas com maior número de colaborações faz sentido, afinal são áreas de excelência no Brasil desde o surgimento das primeiras instituições, como o Instituto Oswaldo Cruz e o Instituto Butantan, que praticamente fundaram a ciência brasileira”, diz ela. As áreas de saúde e biológicas também têm custos altos e por isso, diz Jacqueline, parte das parcerias entre pesquisadores busca compartilhar infraestrutura.

Para entender o alcance dos resultados da pesquisa, é preciso compreender certas limitações dos dados primários. Por exemplo, a distinção que deve ser feita entre colaboração envolvendo indivíduos e aquela celebrada por instituições. “O estudo mostra um aumento da colabora-



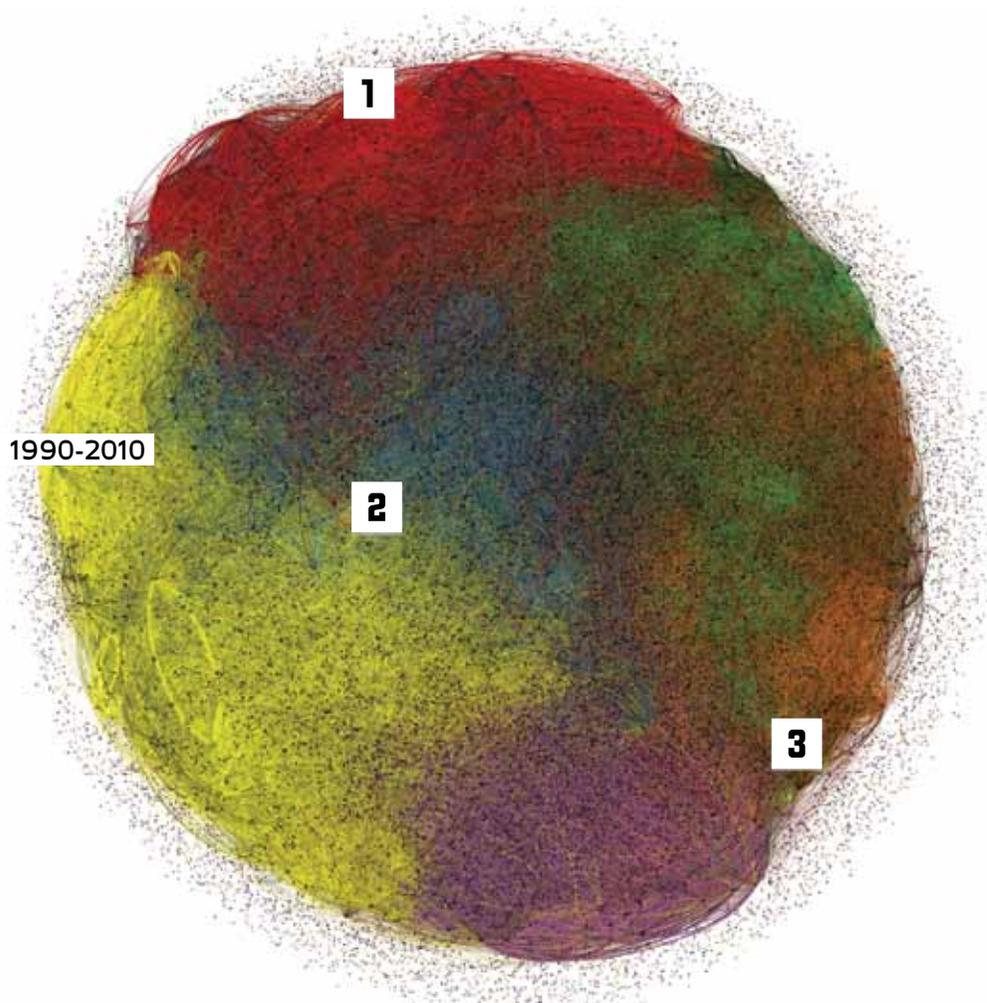
2008-2010

ção entre pessoas, mas isso não significa necessariamente que as instituições de pesquisa estão fazendo mais colaborações”, ressalta Samile Vanz. Segundo ela, a colaboração entre instituições é especialmente importante, porque exige um maior compartilhamento de equipes, de recursos e equipamento. “O país está estabilizando o nível de coautorias individuais e talvez esteja passando por uma nova fase de ampliação da colaboração entre instituições, mas isso só estudos futuros irão mostrar”, diz a pesquisadora, que defendeu em 2009 uma tese de doutorado sobre a participação da pesquisa brasileira em redes internacionais (ver Pesquisa FAPESP n° 169).

RESSALVAS

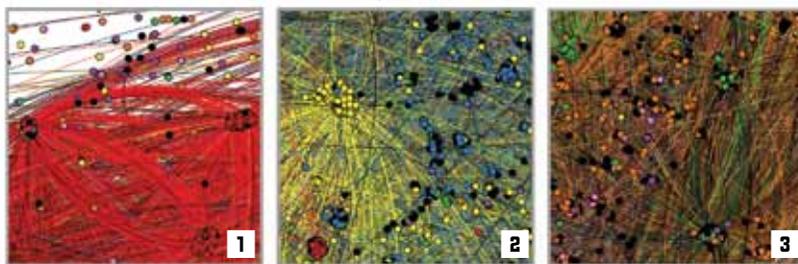
Mena-Chalco explica que o cruzamento de dados concentrou-se na identificação de coautoria entre pesquisadores com registro na Plataforma Lattes, sem levar em consideração a origem institucional deles. Um exemplo: a colaboração entre dois brasileiros com currículo Lattes, sendo um da USP e outro de Harvard, caracteriza-se como uma colaboração internacional. “Mas como a pesquisa leva em conta apenas os indivíduos, não é possível distinguir, entre os resultados, o que é colaboração internacional e o que é nacional. Nosso objetivo foi ver como pesquisadores brasileiros estão se relacionando”, explica o professor da UFABC. Outra limitação é em relação à falta de atualização dos currículos. Se dois pesquisadores escreveram um artigo juntos, mas apenas um deles registrou isso no currículo Lattes, a falta de informação do outro autor impossibilitou a identificação da colaboração. ■

1990-2010



Rede de interações

Cada gráfico em forma de novelo (*acima*) representa uma rede de coautorias entre pesquisadores obtida a partir da Plataforma Lattes em momentos diferentes no período entre 1990 e 2010. Cada pesquisador é representado por um ponto colorido, e a coautoria entre dois pesquisadores, por uma linha colorida cuja composição é a mistura das cores dos pontos que ela liga. Os pesquisadores de cada área estão representados com uma mesma cor (ver *esquema de cores na página ao lado*). Para cada triênio foram obtidas as redes de coautoria considerando exclusivamente produções publicadas nesse período. A divisão por triênios mostra a evolução das coautorias e revela o quanto as interações entre pesquisadores adensaram-se ao longo do tempo.



Os três quadros em destaque mostram detalhadamente como ocorrem as interações dentro de cada área e também entre áreas distintas. Por exemplo, pontos azuis, que representam pesquisadores das ciências biológicas, aparecem na região amarela, onde se concentram pesquisadores da área de ciências da saúde. Isso indica uma intensa interação entre pesquisadores das duas áreas.