

A Pesquisa no Brasil: Promovendo a excelência

Análise preparada para a CAPES
pelo Grupo *Web of Science*

Conteúdo

5

Como é realizada a pesquisa no Brasil?

8

Comparações Internacionais

18

Identificando a excelência na pesquisa Brasileira

30

Apêndice 1

40

Apêndice 2

Esta análise foi preparada para a CAPES pelo Grupo *Web of Science*.

Ele fornece uma visão geral de alto nível do desempenho da pesquisa brasileira, descrevendo tendências recentes.

Agradecemos ao professor Carlos Henrique de Brito Cruz, diretor científico da Fundação de Amparo a Pesquisa de São Paulo (FAPESP) e ao Sr. Abel Packer, diretor e co-fundador da Scientific Electronic Library Online (SciELO) por suas contribuições.

Sumário Executivo

Essa análise fornece um resumo da pesquisa brasileira em um contexto global e destaca a importância das colaborações internacionais e com a indústria na visibilidade e no impacto da pesquisa. Usando a bibliometria para analisar trabalhos brasileiros de pesquisa publicados entre 2013 e 2018, exemplos ilustrativos de análises e evidências descrevem o cenário da pesquisa e fornecem uma indicação do impacto potencial das políticas brasileiras de pesquisa e ciência. Os dados são extraídos da *Web of Science* - o maior e mais confiável índice de citações, neutro e independente de editores, do mundo.

Produtividade

O Brasil ocupa a 13ª posição no mundo em termos de produção de artigos e revisões de pesquisa indexados na *Web of Science*.

Somente em 2018, pesquisadores brasileiros publicaram mais de 50.000 artigos.

O crescimento da produção é de 30% nesse período de seis anos observado e é o dobro da média global.

Existem bolsões de excelência, em termos de impacto de citações, na pesquisa brasileira, nas ciências da vida, ciências físicas e engenharia.

Impacto da Citação

A porcentagem de papéis brasileiros no primeiro por cento dos artigos mais citados do mundo é consistentemente próximo a um por cento durante o período de seis anos, 2013-2018.

O Impacto de Citação Normalizado por Categoria (CNCI) desta pesquisa está abaixo da média mundial, mas melhorando constantemente.

Comparação entre pares

Comparado a outros países do BRICS, a produção brasileira de pesquisa, em termos do número de artigos indexados na *Web of Science* é média.

Comparado aos países vizinhos da América Latina, a produção brasileira de pesquisa tem um CNCI mais baixo.

Colaboração Internacional:

Entre 2013 e 2018, pesquisadores brasileiros colaboraram com pesquisadores de 205 países.

Aproximadamente um terço dos trabalhos de pesquisa do Brasil são de coautoria com pesquisadores de outros países.

O crescimento da porcentagem de artigos com um coautor internacional diminuiu de 17,5% entre 2013 e 2015 para um aumento de 1,8% entre 2016 e 2018.

Colaboração com Indústria

Estratégias para promover colaborações de pesquisa entre a universidade e a indústria estão dando frutos em termos do número de artigos publicados.

81% das publicações conjuntas da universidade e indústria para o período 2015-2017 são colaborações entre universidades públicas e indústria.

Universidades e Institutos de Pesquisa

As universidades públicas são a principal fonte de publicações de pesquisa no Brasil.

As 15 universidades com maior produção de pesquisa, todas públicas, produzem mais de 60% da produção total de pesquisa.

Histórico

Na última década, o Brasil registrou um crescimento considerável em sua pesquisa, mantendo seu status como um dos países do BRICS, ao lado da Rússia, Índia, China e África do Sul, e como uma economia emergente, demonstrando um rápido crescimento da produtividade e influência científica medidas através da bibliometria. O Brasil é o quinto país mais populoso do mundo e permanece entre as 10 principais economias do mundo, apesar da recente turbulência econômica.

Nesta análise atualizada, o perfil de pesquisa do Brasil em 2017 é revisitado¹ para examinar onde os pesquisadores brasileiros continuam ativos e para identificar áreas de excelência. Na última década, houve um interesse crescente na produção de pesquisa do Brasil focada no nível nacional² e em campos específicos de pesquisa.³ Observando o cenário de pesquisa em rápida mudança no Brasil, Leta, Thijs e Glänzel (2013) observam que “ O enorme crescimento da produção brasileira de publicações constitui o maior potencial que vai muito além da região da América Latina”.

O relatório de 2017 documentou o estado da produção da pesquisa indexada na *Web of Science*. Nesta análise, exploramos os dados da *Web of Science*⁴, com foco nas colaborações internacionais e colaborações universidade - indústria e na produção de pesquisas de universidades e institutos de pesquisa para oferecer exemplos ilustrativos de como diferentes análises podem gerar evidências para ajudar a informar políticas públicas e tomadas de decisão.

Também examinamos alguns resultados do Scientific Electronic Library Online (SciELO), que foi criado e implementado no Brasil em 1998 para fortalecer a infraestrutura de pesquisa, desenvolvendo a capacidade de comunicação acadêmica. Atualmente, o SciELO inclui 296 periódicos brasileiros, dos quais 99 são periódicos influentes de alta qualidade indexados na *Web of Science*.

Ao estudar a produção científica de um país como o Brasil, é aconselhável focar a análise bibliométrica em documentos da *Web of Science* classificados como artigos ou revisões, pois são publicações revisadas por pares que refletem resultados de pesquisas maduras.

¹ <https://www.capes.gov.br/images/stories/download/diversos/17012018-CAPES-InCitesReport-Final.pdf> acessado em 1 de agosto de 2019. Adams J. and King C. (2009) Global Research Report: Brazil - Research and Collaboration in the New Geography of Science. <https://www.slideshare.net/nielsleidecker/grr-brazil-jun09-1>

² Glänzel, W., Leta, J. Thijs, B. (2006) Science in Brazil Part 1: A macro-level comparative study. *Scientometrics*, 67 (1), 67–86. Leta, J, Thijs, B, & Glänzel, W. (2013) A macro level study of science in Brazil: seven years later. *Bibli Encounters: Electronic Journal of Library and Information Science* [Online], acessado em 25 de Agosto de 2019

³ Leta, J., Glänzel, W., Thijs, B. (2006), Science in Brazil. Part 2: Sectorial and institutional research profile. *Scientometrics*, 61 (1), 87-105. Andrade Vargas, R.; de Souza Vanz, S.A; Chittó Stumpf, I.R. (2015) Brazilian agricultural research in the *Web of Science*: a bibliometric study of scientific output and collaboration (2000-2011). *Em Questão*, 21 (3), 296-318.

⁴ Ao longo deste documento, a análise é restrita a artigos e revisões, indexados no *Science Citation Index Expanded* (SCIE), o *Social Science Citation Index* (SSCI) e o *Arts and Humanities Citation Index* (AHCI) da *Web of Science*. Sempre que mencionado, análises adicionais são baseadas na coleção brasileira do *SciELO Citation Index*, hospedada na plataforma *Web of Science*.



Como é realizada a pesquisa no Brasil?

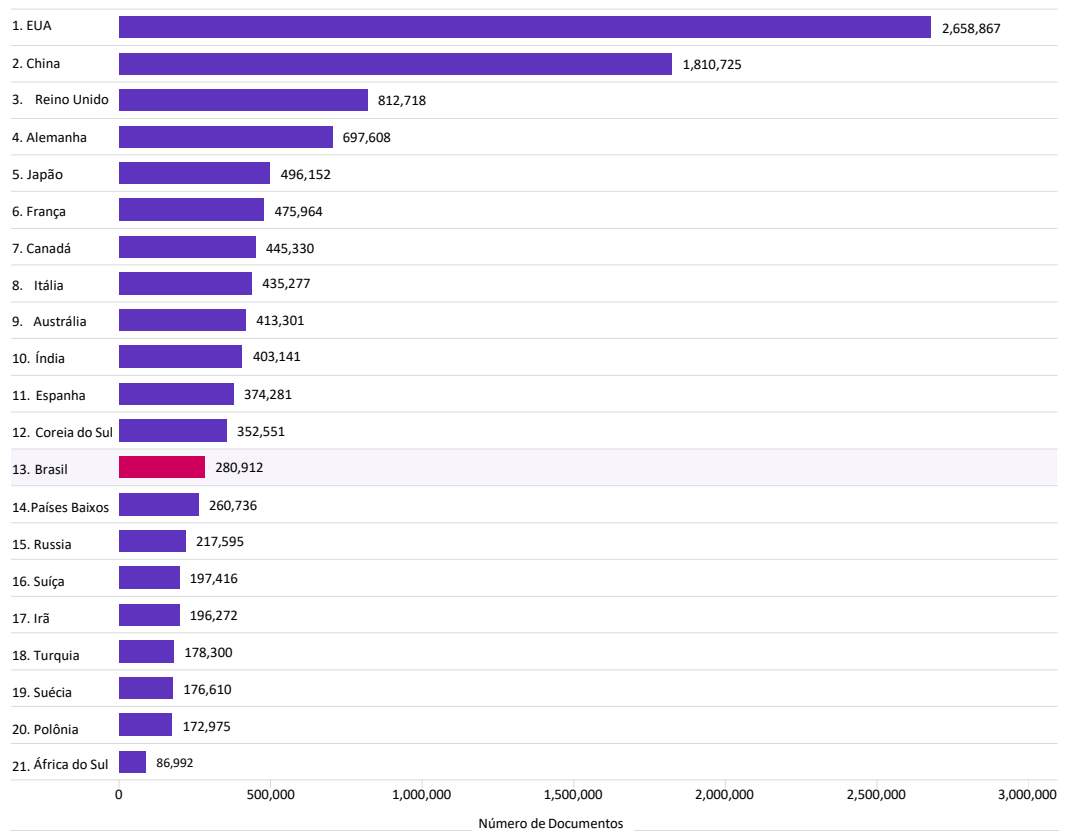
Uma visão geral do desempenho da pesquisa brasileira nos últimos anos

Quantos artigos científicos o Brasil produz?

O Brasil ficou em 13º lugar no mundo em termos de produção de trabalhos de pesquisa entre 2013 - 2018, logo atrás da Índia (10º) e Coreia do Sul

(12º), e à frente da Rússia (15º) e África do Sul (21º) (Figura 1). Vale ressaltar que, apesar das mudanças nas condições econômicas, o crescimento da produção brasileira de pesquisa tem sido consistentemente forte ao longo deste período de seis anos.

Figura 1



Artigos e Revisões indexados na *Web of Science*, 2013 – 2018.

O crescimento de 30% no número de publicações nesse período de seis anos foi o dobro da média global (15%). Somente em 2018, pesquisadores brasileiros publicaram mais de 50.000 artigos indexados na *Web of Science*.

Além dos 280.912 trabalhos indexados na *Web of Science* (2013-2018), autores brasileiros publicaram mais de sessenta mil artigos e revisões indexados em periódicos SciELO que não são abordados na *Web of Science*.

Os estudos bibliométricos costumam usar o 1% e o 10% dos artigos mais citados no mundo como indicadores padrões para a produção de pesquisa de excelência. Na Figura 2, ao lado esquerdo, mostra a porcentagem de artigos publicados pelo Brasil que estão no top 1 % dos artigos mais citados

do mundo, e o painel direito mostra a porcentagem de artigos publicados pelo Brasil que estão entre os 10% mais citados do mundo. Ao analisar os 10% dos trabalhos mais citados, o Brasil está abaixo da média mundial e a queda nos últimos anos parece mais proeminente. Essa aparente queda não é exclusiva do Brasil, uma vez que maiores quedas nessa medida são observadas em dados recentes de países que possuem extensos resultados de pesquisa. É importante enfatizar que se deve ter cuidado na interpretação das contagens recentes de citações a partir de 2018, que tendem a ser voláteis e os aparentes desvios da norma podem ser artefatos de recente evolução dos dados, em vez de desvios dramáticos de tendências persistentes.⁵

Figura 2



Percentual de artigos Brasileiros no top 1% e top 10% dos artigos mais citados, 2013 – 2018 (media mundial indicada na linha pontilhada).



Comparações Internacionais

Qual o impacto de citação da pesquisa brasileira?

À medida que a produção da pesquisa brasileira cresceu entre 2013-2016, o impacto de sua pesquisa, como indicado pelo número de citações recebidas, também cresceu de forma constante. A Figura 3 apresenta uma perspectiva global seletiva sobre a distribuição do Impacto de Citação Normalizado da Categoria (CNCI)⁶.

O CNCI anual do Brasil atingiu um pico em 2016 em 0.91, em comparação com uma média mundial de 1.0. No entanto, observa-se uma queda no CNCI do Brasil em 2017.

Outros países da América Latina⁷ também apresentam essa tendência. Essa tendência não é replicada nos países desenvolvidos⁸, assim como também não é replicada no BRICS⁹, onde o aumento proporcional dramático nas publicações e nas publicações de alto impacto da China mascaram os efeitos em outros países.

Essa queda em 2017 não se deve a uma mudança na cobertura de periódicos brasileiros na *Web of Science*. Uma explicação potencial para essa diminuição do CNCI é o reduzido crescimento no número de artigos com um ou mais coautores internacionais.

É amplamente reconhecido que colaborações internacionais melhoram o impacto de citação de um artigo¹⁰. Nos três anos, de 2013 a 2015, o percentual de artigos com coautor internacional aumentou 17.5%, já nos três anos mais recentes de 2016 a 2018 houve apenas 1.8% de crescimento. Porém, como mencionado acima, uma explicação provável dessa aparente desaceleração no CNCI brasileiro é que ele ainda está sujeito à variação dos dados mais recentes

Figura 3



Impacto de Citação Normalizado da Categoria (CNCI) do Brasil¹¹ e países comparativos por grupos, 2013 – 2017.

⁶Veja o anexo 1 para mais detalhes sobre o cálculo do CNCI.

⁷Argentina, Chile, Colômbia, México

⁸Canadá, França, Alemanha, Japão, Espanha

⁹Rússia, Índia, China, África do Sul

¹⁰Katz, J.S. & Hicks, D. *Scientometrics* (1997) 40: 541. <https://doi.org/10.1007/BF02459299>, acesso em 21 de agosto, 2019. Adams, J. & Loach, T. (2015). A well-connected world. *Nature*. 527. 558-559.

¹¹Os diamantes e os números acima dos diamantes indicam o valor do CNCI do Brasil.

Como a pesquisa brasileira se compara internacionalmente?

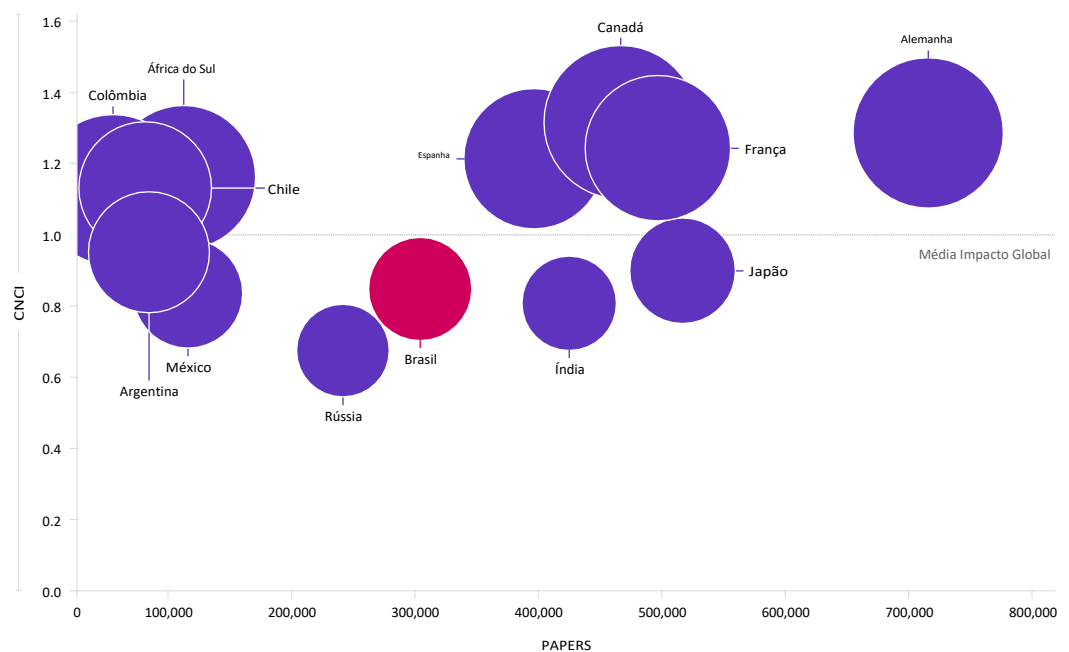
Vários fatores, como, a força da economia, o volume e foco dos financiamentos de pesquisa, as colaborações internacionais e o calibre das instituições de ensino e pesquisa, podem afetar a quantidade e influência da produção de pesquisa de um país. Diferentes países variam em suas abordagens à políticas de ciência e tecnologia, portanto, embora seja difícil expor a contribuição de cada um desses diferentes fatores, pode ser mais esclarecedor quando se compara a produtividade e o impacto da pesquisa

entre os países para obter maior clareza sobre a relação entre esses fatores e os resultados da pesquisa.

Por exemplo, os recentes investimentos da China em ensino superior e pesquisa, a levaram a um crescimento exponencial na quantidade e qualidade da produção científica, medida pelo número de citações.

Nesta análise do Brasil com seus países comparativos, a China, que publicou pelo menos três vezes mais artigos do que o Brasil ou qualquer outro país comparativo, é um extremo que distorce a visualização dos dados, logo, os dados da China são incluídos na análise, mas excluídos da visualização Figura 4.

Figura 4



Produção e Impacto de Citação Normalizado da Categoria (CNCI) de todos os países comparativos excluindo China. O diâmetro do círculo é proporcional ao percentual de documentos (Artigos e Revisões entre os 1% dos artigos mais citados do mundo, 2013 - 2018).

Em comparação com os outros países do BRICS, a produção de pesquisa do Brasil é média, com mais publicações do que as da África do Sul e Rússia e um impacto de citação maior do que a Rússia e Índia.

A China tem maior impacto de citação, está acima da média mundial, e tem a maior produção de artigos revisados por pares.

Comparado aos países desenvolvidos, o Brasil tem um menor impacto de citação. O impacto da pesquisa no Brasil está apenas ligeiramente atrás do Japão, que é o único país comparador desenvolvido com um impacto de citação abaixo da média mundial (Figuras 3, na página 9 e Figura 4, na página 10).

A Figura 4 também mostra que o Brasil publica muito mais artigos do que outros países da América Latina, embora os outros países tenham maior impacto de citação e proporção de artigos em 1% daqueles mais citados mundo. O maior impacto de citação é provavelmente derivado da colaboração internacional. Chile e Colômbia produzem a maior parte de suas pesquisas, 67% e 67,5%, respectivamente, em colaboração com pesquisadores de outras nações, enquanto esse percentual para o Brasil é de 36% no mesmo período. Análises adicionais de tendências forneceriam informações úteis sobre a posição relativa do Brasil e como ela evoluiu ao longo do tempo em relação a outros países.

Colaboração internacional e seu impacto de citação

Com quem o Brasil colabora internacionalmente?

Pesquisadores brasileiros colaboram com acadêmicos de todo o mundo. No período de seis anos, 2013-2018, brasileiros foram coautores de trabalhos de pesquisadores de 205 países, o que representou aproximadamente um terço de todos os trabalhos da *Web of Science* com autores brasileiros. Essas colaborações incluíram países reconhecidos por suas intensas pesquisas, como o G7, países vizinhos da América Latina e outros países do BRICS. Como mencionado anteriormente, as colaborações internacionais geralmente têm um impacto maior na citação.

O maior número de artigos em coautoria internacional é publicado com autores nos Estados Unidos, embora esses artigos não tenham o maior impacto de citação. Alguns dos mais altos impactos de citação estão associados a publicações de parceiros em outros países do BRICS, como China e Índia, onde a produção de pesquisas vem crescendo rapidamente nos últimos anos.

Novamente, uma análise mais detalhada dos trabalhos com múltiplos coautores, sugere uma diferença qualitativa entre colaborações de trabalho realizado em instituições de pesquisa hiper-colaborativas, como o CERN, onde trabalhos com centenas de coautores recebem milhares de citações *versus* artigos que são o resultado de pequenos grupos de pesquisadores trabalhando em colaboração em projetos relativamente pequenos.

O Brasil também está desenvolvendo mais colaborações regionais na América Latina, conforme observado no Relatório Global de Pesquisa sobre o Brasil (Adams & King, 2009)¹². Os investimentos em projetos como o Laboratório Nacional de Luz Síncrotron (LNLS), a única fonte de luz síncrotron na América Latina, melhoram oportunidades para colaborações internacionais.

¹² Adams J. and King C. (2009) Global Research Report: Brazil - Research and Collaboration in the New Geography of Science. <https://www.slideshare.net/nielsleidecker/grr-brazil-jun09-1>

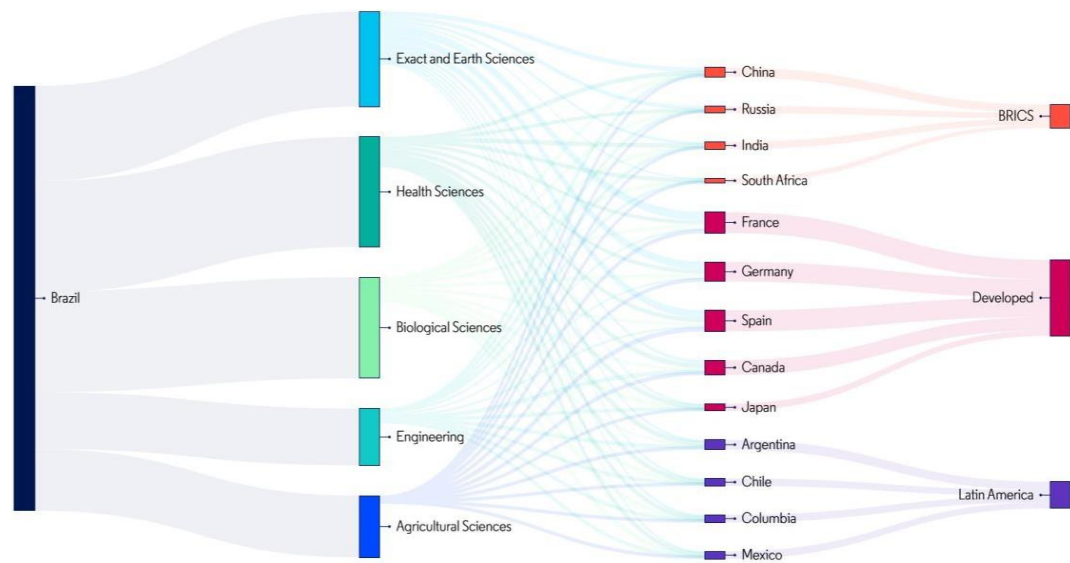
Este laboratório abriga cerca de 1.200 pesquisadores brasileiros e estrangeiros envolvidos em centenas de projetos e publicações em coautoria¹³.

Análises detalhadas da produção brasileira de pesquisas com colaborações internacionais são apresentadas em nosso relatório CAPES 2017¹⁴, para um conjunto de países seletos. Uma análise e visualização alternativas dessas colaborações internacionais são demonstradas na Figura 5, mostrando a proporção de artigos publicados entre 2013 e 2018 em um subconjunto selecionado das nove categorias de pesquisa da CAPES relativas a coautores internacionais.

A barra à esquerda representa todos os trabalhos publicados neste período de seis anos, em cinco das nove categorias de pesquisa da CAPES com o maior número de publicações.

Linguística, Literatura e Artes, Ciências Humanas e Sociais, Ciências Sociais Aplicadas e Multidisciplinar, com menor número de publicações, estão excluídas desta análise. A distribuição do número de artigos em cada uma das cinco categorias restantes é mostrada pelo tamanho das barras no meio. Os fluxos das barras do meio para os países da direita são proporcionais ao número de artigos com coautores internacionais dos países do lado direito. Há uma contagem dupla nesses fluxos, porque vários trabalhos têm coautores de vários países. O tamanho dos retângulos é proporcional ao número de artigos em coautoria, mostrando a Espanha com os trabalhos com mais coautoria, o que é notável por ter uma produção geral menor de artigos em comparação com outros países desenvolvidos e alguns países do BRICS.

Figura 5



Colaboração internacional com países selecionados ¹⁴ por diferentes áreas de pesquisa

¹³ <https://www.inls.cnpem.br/the-inls/about/> - acesso em 17 de agosto, 2019.

¹⁴ <https://www.capes.gov.br/images/stories/download/diversos/17012018-CAPES-InCitesReport-Final.pdf> - acesso em 17 de agosto, 2019.

Entre esses países comparativos, os países desenvolvidos têm o maior número de colaborações com o Brasil. No entanto, o número de artigos colaborativos não parece proporcional à produção total de publicação do país parceiro. Se assim fosse, seria de esperar que a China tivesse o maior número de papéis colaborativos. Nota-se que o número de colaborações depende das redes de pesquisa existentes e com pesquisadores de países parceiros com um histórico de excelência em ciências.

Proporcionalmente, o menor número de colaborações é em Ciências Agrícolas e o maior em Ciências Exatas e da Terra. Para cada país parceiro, as Ciências Exatas e da Terra é geralmente a área de pesquisa mais produtiva, com pelo menos 1.000 artigos por parceria. O número mais alto de artigos está em países europeus, com a França colaborando em 6.381 artigos em Ciências Exatas e da Terra, que é o maior número de artigos em uma única disciplina.

Uma grande proporção de colaborações em Ciências da Saúde e Ciências Biológicas é relacionada aos EUA (40 a 50%), seguido por países europeus e Argentina, com os países do BRICS desempenhando um papel muito menor nessa colaboração.

Em todas as áreas de pesquisa, o Japão tem o menor número de artigos em colaboração com o Brasil. Esses números baixos são ligados com a menor porcentagem geral de artigos do Japão com coautores internacionais (31,8%) em comparação com outros parceiros desenvolvidos que possuem 50% ou mais.

Os países da América Latina têm a menor quantidade de documentos colaborativos em comparação com os países parceiros do Brasil de outras partes do mundo. Argentina colabora com mais frequência com o Brasil em Ciências Agrárias (439 artigos) e Ciências Biológicas (1.762). Em toda a região, a maioria das colaborações estão em Ciências Exatas e da Terra e Ciências da Saúde e a minoria em Engenharia.

No SciELO, as colaborações internacionais cresceram constantemente nos últimos anos e representam quase 10% de todas as publicações em 2018. As colaborações mais frequentes, indexadas no SciELO, são com pesquisadores dos EUA, Portugal, China, Argentina, Espanha e Turquia.

Os países do BRICS colaboram com pouca frequência em Ciências Agrícolas e Ciências Biológicas.

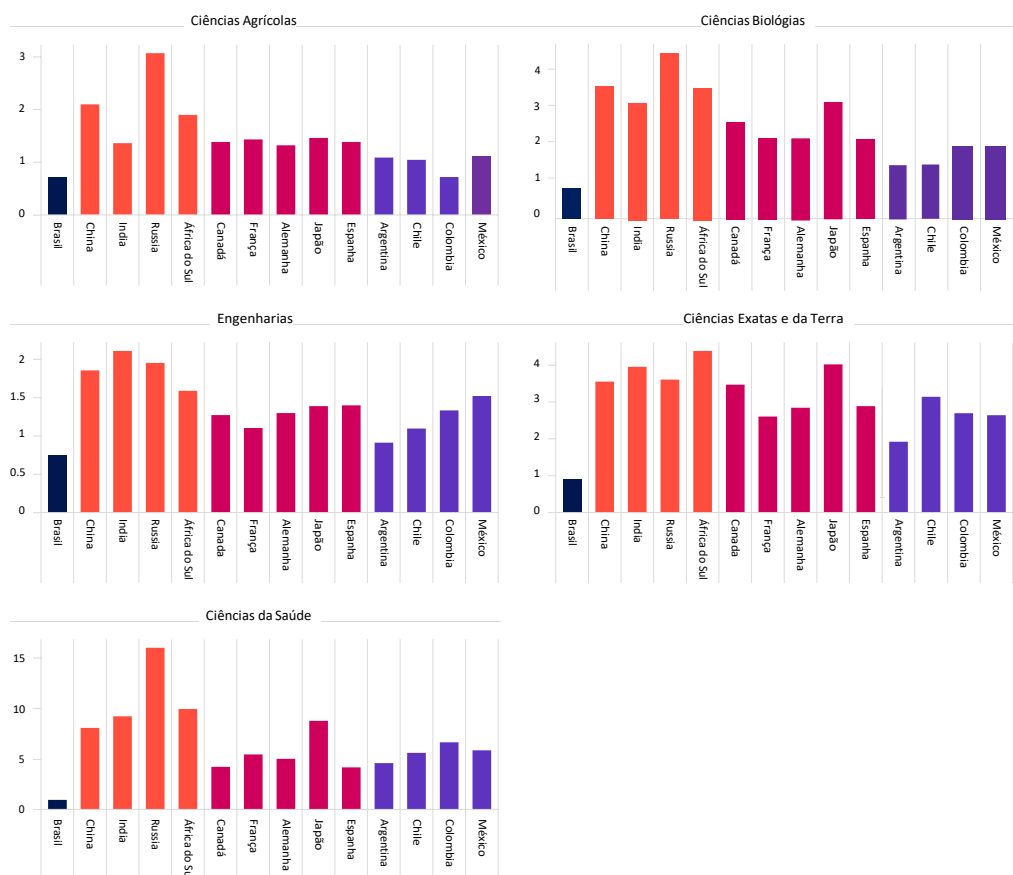
Há uma maior colaboração da China e da Rússia em Ciências Exatas e da Terra e com a China e a Índia nas Ciências da Saúde.

Qual o impacto de citações das colaborações internacionais?

O impacto de citações dos trabalhos com coautores internacionais é geralmente mais alto que o de trabalhos publicados somente por autores do Brasil.¹⁵ Para publicações de autores Brasileiros e coautores internacionais, este número é consistentemente alto (Figura 6), e acima da média mundial de 1.0. As poucas exceções onde o CNCI (impacto de citações normalizado da categoria) está abaixo de 1.0 são para as colaborações com países da América Latina em Ciências Agrícolas

(Colômbia, 0.72) e Engenharia (Argentina, 0.91). Os índices de CNCI mais baixos são em Ciências Agrícolas onde eles variam entre 0.72 e 3.17 (Rússia). Mesmo em Ciências Agrícolas, o CNCI é extremamente alto para colaboração com os BRICS. Trabalhos publicados em Ciências da Saúde tem o mais alto impacto de citações, variando de 4.13 (Espanha) até o valor extremamente alto de 15.51 (Rússia).

Figura 6



Impacto de Citações Normalizado da Categoria (CNCI) de publicações Brasileiras com coautores de países selecionados por área do conhecimento

Esses altos índices de impacto de citações devem ser analisados com critério. As publicações que possuem esses índices incluem vários trabalhos hiper-colaborativos com muitos autores e muitos países afiliados tratando de tópicos como experimentos de física de partículas, astronomia, guias médicas ou sobre

impacto global de doenças. Tais trabalhos são frequentemente citados milhares de vezes. O impacto de citações do Brasil tende não a ser devido a colaborações regionais; mas concentrado em trabalhos com coautores de países de outras regiões, como EUA, Europa, e outras partes do mundo.

¹⁵ A relação entre coautoria e impacto de citações é um fenômeno que tem sido amplamente observado e é discutido em Moed, H. (2005) Citation Analysis in Research Evaluation. Dordrecht: Springer pp285-290.

Colaboração Universidade-Indústria

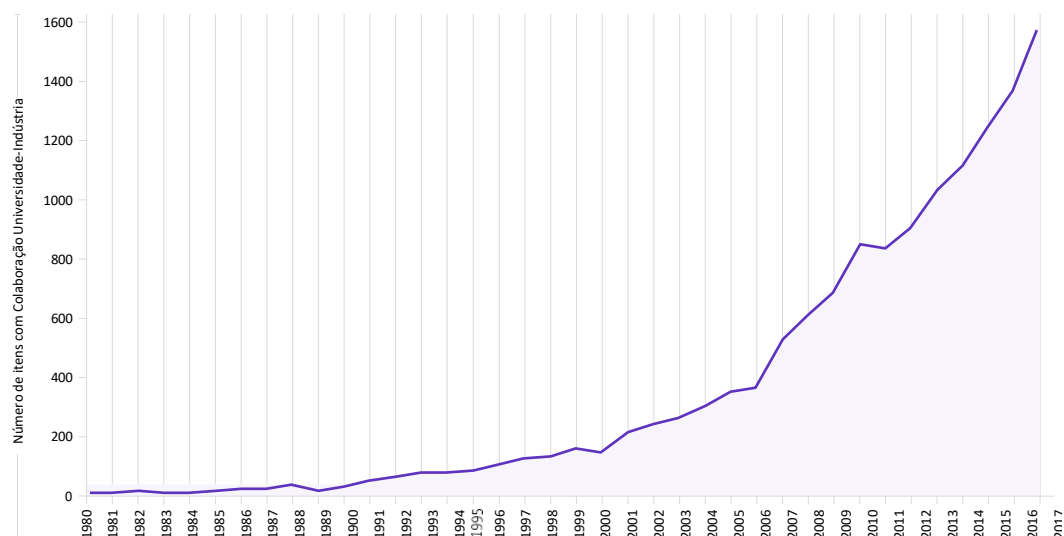
Como a colaboração na pesquisa entre a universidade e a indústria tem se modificado em 37 anos?

A interação entre universidade e indústria tem estado como objetivo central de Políticas Brasileiras de Ciência e Tecnologia por muitas décadas. Paradoxalmente, muito pouca análise e medições de quantidade e intensidade desta interação tem sido feitas. Aqui contribuimos com uma análise baseada em números das publicações científicas e técnicas que possuem coautoria entre pesquisadores de universidades no Brasil e pesquisadores na indústria¹⁶ (para a discussão da metodologia, veja o Apêndice 2). Com o intuito de prover um contexto mais amplo para essas colaborações universidade-indústria, uma visão de mais longo termo, iniciando-se em 1980 é mais adequada, preferencialmente a usar o período de seis anos utilizado no restante desta análise.

O número de publicações com coautoria entre pesquisadores em universidades e na indústria oferece uma visão das ideias que foram em conjunto criadas e desenvolvidas por esses pesquisadores nesses dois setores indicando um nível mais elevado de engajamento que meramente consultivo, contratação de pesquisa e desenvolvimento, ou suporte à pesquisa e doações. A Figura 7 mostra a evolução dessas interações de pesquisa desde 1980, pela contagem do número de itens na *Web of Science* (todas as categorias) que possuem pelo menos um autor de uma universidade no Brasil e um coautor da indústria, de qualquer lugar do mundo.

A Figura 8 mostra a evolução da coautoria entre universidades no Brasil e autores do setor industrial desde 1980 para as dez universidades com o maior número de publicações conjuntas. As dez universidades mostradas – todas públicas – são responsáveis por 81% das publicações conjuntas para o período 2015-2017.

Figura 7

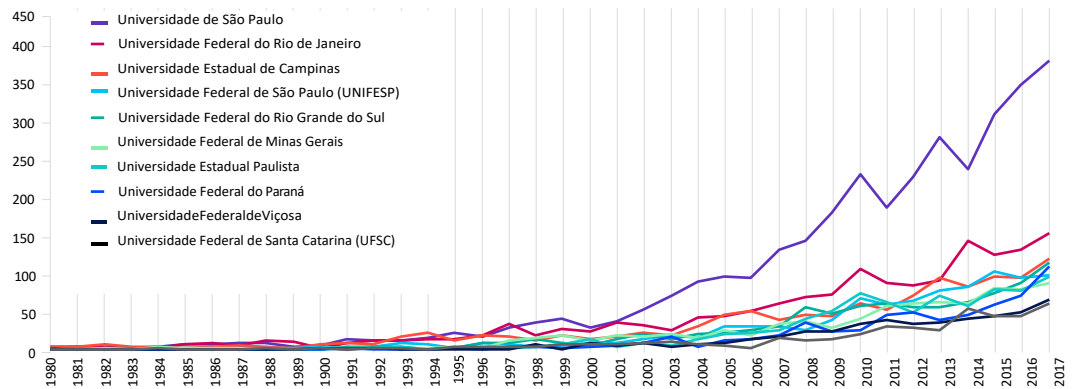


Número de itens na *Web of Science* com pelo menos um autor em Universidade no Brasil e pelo menos um coautor da indústria.

¹⁶ Os dados aqui apresentados da colaboração universidade-indústria incluem um conjunto de organizações corporativas muito maior que os atualmente considerados na *Web of Science* ou Incites. Esses dados foram obtidos fazendo-se um procedimento de busca avançada que incluiu a reclassificação da natureza das organizações na lista de publicações com autores no Brasil e também, pesquisando-se por entidades que possuam associação com a indústria.

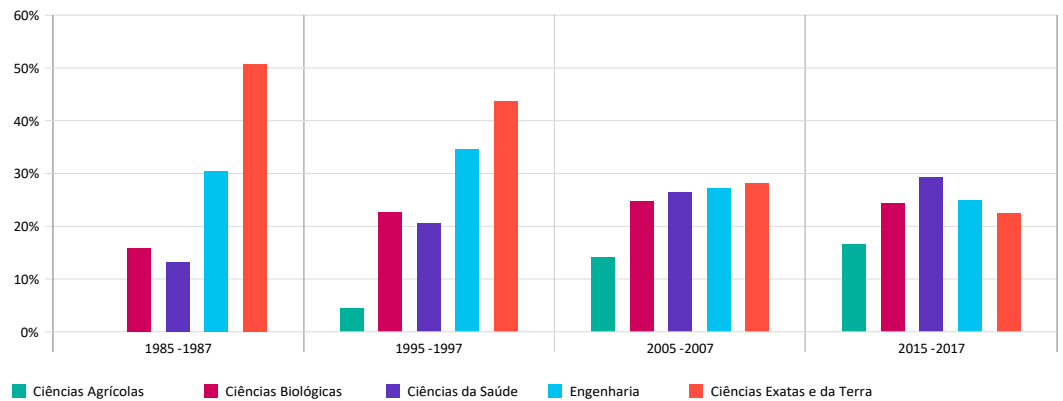
Quais universidades possuem mais coautoria com a indústria?

Figura 8



Número de itens na *Web of Science* com pelo menos um autor em uma universidade no Brasil e pelo menos um coautor da indústria para as dez universidades com as maiores quantidades de itens em 2017.

Figura 9



Classificação de publicações em coautoria entre universidades no Brasil e companhias de acordo com as "Áreas de Pesquisa" definidas pela CAPES.

A distribuição das publicações em coautoria universidade-indústria de cinco das nove áreas de pesquisa definidas pela CAPES está mostrada na Figura 9, para quatro períodos de três-anos: 1985-87, 1995-97, 2005-07, e 2015-17. Somente cinco das nove categorias CAPES estão mostradas aqui devido ao pequeno número de colaborações nas categorias de Linguística, Literatura e

Artes, Humanidades e Ciências Sociais, Ciências Sociais Aplicadas e Multidisciplinar. A distribuição que estava fortemente dominada pela área de Ciências Exatas e da Terra há 30 anos atrás, tem se distribuído ao longo do tempo entre as cinco categorias, especialmente, entre as Ciências da Vida e Físicas para o período de 2015-17.

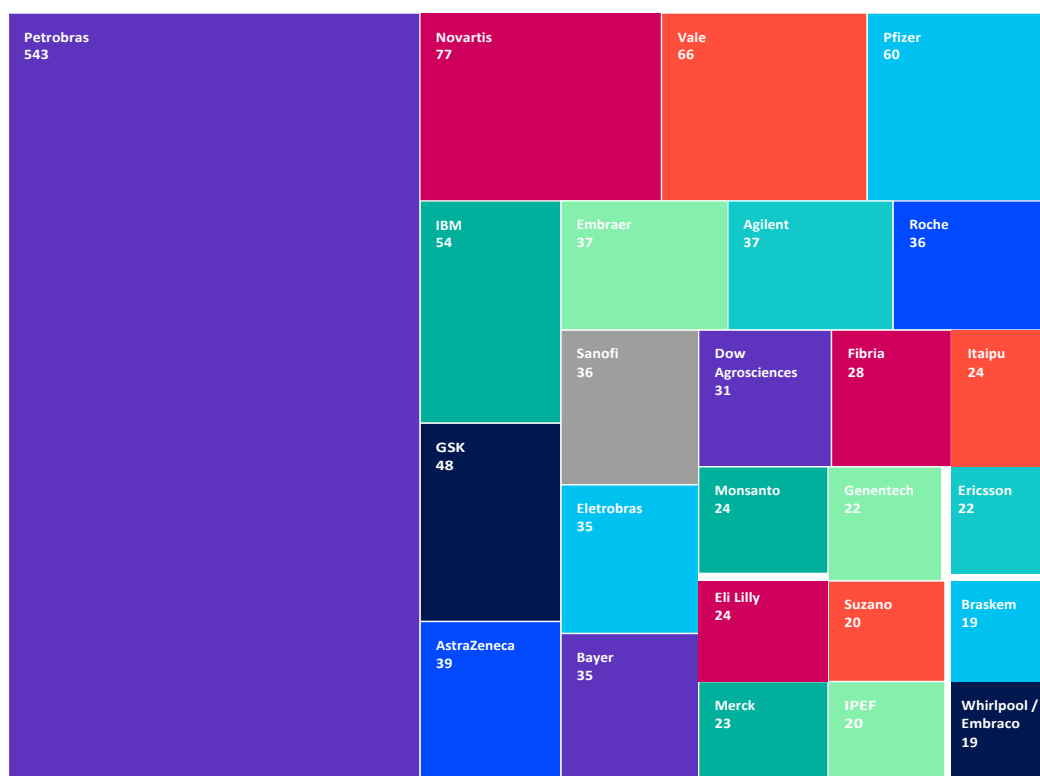
Quais companhias publicam em colaboração com universidades no Brasil?

14%

Petrobrás participou em 14% de toda a produção que teve colaboração universidade-indústria.

Petrobrás, que possui um extenso programa para o desenvolvimento de colaborações de pesquisa com a academia, domina o panorama da colaboração da universidade-indústria no Brasil (Figura 10). No período 2015-2017, com 543 publicações, a empresa participou em 14% de toda a produção de colaboração entre universidade e indústria.

Figura 10



As 25 empresas que tiveram as maiores quantidades de publicações em coautoria com universidades no Brasil entre 2015-2017.

Das 50 companhias com os maiores níveis de coautoria com a academia, 17 são brasileiras e 33 são multinacionais. Há uma predominância de companhias do setor farmacêutico (18 de 50), seguidas por empresas do setor agrícola (12 de 50).

A criação de Fundos do Setor de Ciência e Tecnologia, para financiar a pesquisa, desenvolvimento, e inovação do país tem estimulado desenvolvimentos de economia sustentável e colaboração de pesquisa universidade-indústria na área de bio-etanol e bio-farma.¹⁷

¹⁷ Financial Times Confidential Research, (2015) Brazil – Biopharma seeking a pick-me-up, <http://www.scienceforbrazil.com/wp-content/uploads/2015/12/Brazil-%E2%80%93-Biopharma-seeking-a-pick-me-up.pdf> accessed August 22, 2019. Cortez, LAB, Nogueirab, LAH, Lealc, MRVL, Junior, RB. 40 Years of the Brazilian Ethanol Program (Proálcool): Relevant Public Policies and Events Throughout Its Trajectory and Future Perspectives http://bioenfapesp.org/gsb/lacaf/documents/papers/05_ISAF_2016_Cortez_et_al.pdf accessed August 22, 2019.



Identificando excelência na Pesquisa Brasileira

Como se compara a pesquisa Brasileira em diferentes áreas?

Para focalizar na excelência da pesquisa, entre 2013-2018, a análise das nove categorias definidas pela CAPES nos oferece grande visibilidade. A excelência nesse contexto é medida pelo Impacto de Citações Normalizado da Categoria (CNCI).

Entre essas nove categorias CAPES, a maioria dos trabalhos de pesquisadores Brasileiros são publicados nas áreas de Ciências da Vida e Físicas e Engenharia. O meio principal para disseminação de resultados de pesquisa nessas áreas é a publicação em periódicos, que são bem representados na base da *Web of Science*. Por outro lado, em artes, humanidades e ciências sociais, os pesquisadores mais frequentemente escrevem livros ou documentos normativos, os quais são menos representados nesta base.

Devido em parte aos seus baixos números na *Web of Science*, a produção da pesquisa brasileira em Ciências Sociais Aplicadas e Ciências Humanas possui um alto impacto de citações, acima da média mundial, (Figura 11). Porque o CNCI é uma média ponderada, os trabalhos altamente citados nas sub-categorias de Teologia (145 trabalhos com CNCI médio de 2.37), Arqueologia (224 trabalhos com CNCI 1.24) e Antropologia (393 trabalhos com CNCI de 1.36) elevam o impacto total da pesquisa acima da referência mundial de

1.0. Teologia é notável em outro aspecto em que somente 18.6% dos trabalhos possuem um coautor internacional e, no entanto, seu CNCI é o maior entre as 121 sub-categorias CAPES. Assim, a excelência é baseada na qualidade doméstica na área.

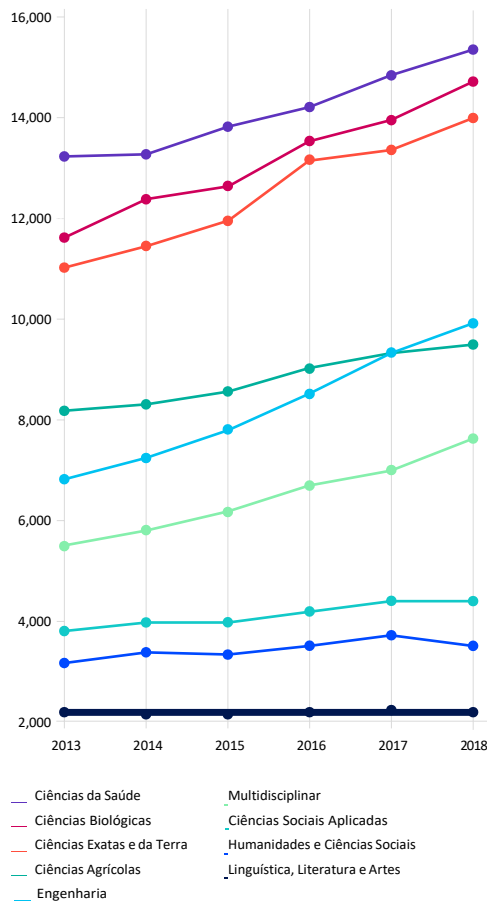
Como mostrado na Figura 11, mais de um terço dos trabalhos em Ciências Biológicas e quase a metade em Ciências Exatas e da Terra possuem colaboradores estrangeiros, o que, como previamente observado, contribui para CNCI mais elevados.

Figura 11

Área do Conhecimento	Trabalhos	CNCI	% Colaboração Internacional
Ciências da Saúde	82,406	0.96	34.7
Ciências Biológicas	75,717	0.74	37.0
Ciências Exatas e Terra	71,214	0.90	45.1
Ciências Agrícolas	46,222	0.71	21.7
Engenharia	42,506	0.76	37.1
Multidisciplinar	30,190	0.82	39.5
Ciências Sociais Aplicadas	14, 29	1.03	30.1
Humanidades e Ciências Sociais	9, 81	1.00	30.5
Linguística, Literatura e Artes	953	0.68	19.1

Produção e Impacto de Citações Normalizado da Categoria (CNCI) de trabalhos Brasileiros publicados entre 2013 e 2018 nas nove categorias de pesquisa CAPES

Figura 12



Tendências em número de trabalhos publicados anualmente em categorias de pesquisa CAPES entre 2013 e 2018.

O foco da pesquisa Brasileira tem evoluído nesses seis anos (Figura 12). Existe um aumento na produção em todas as áreas de pesquisa com exceção de Linguística, Literatura e Artes. Os menores crescimentos foram em Ciências da Saúde (18.7%) e Ciências Agrícolas (21.9%). Embora não sejam tão baixos, esses incrementos parecem pequenos comparados à Engenharia, que cresceu em quase dois-terços, passando a produção de Ciências Agrícolas em 2018.

Se nos aprofundarmos nas 121 sub-categorias CAPES, teremos insights de como seu crescimento tem acontecido. Por um longo período em Engenharia, Ciências de Materiais e Metalurgia tem sido uma área altamente produtiva que tem experimentado um rápido crescimento nos anos recentes. Há também um crescimento considerável em áreas intimamente relacionadas em Ciências Exatas e da Terra (Química, Física, etc). O amplo conhecimento e especialização do Brasil nessas duas áreas relacionadas, e a pesquisa multidisciplinar que as une parecem ter contribuído para esse rápido crescimento.

A análise bibliométrica pode ser usada para determinar áreas de pesquisa que demonstram importância, fraquezas, potencial para oportunidades, ou áreas que estão ameaçadas dentro do portfólio da pesquisa. A Figura 13 mapeia a pesquisa do Brasil nas nove categorias CAPES comparadas com índices comparativos da produção global e impacto de citações.

O eixo horizontal mede a produção proporcional relativa que é a relação representando a proporção da produção de pesquisa do país relativa à proporção da área de pesquisa na produção mundial. Por exemplo, se a parte de Biologia da produção de pesquisa no Brasil é 4.5% e a representatividade da produção global de Biologia indexada na *Web of Science* é 1.5%, então o valor para Biologia no eixo horizontal seria 3. O eixo vertical é o Impacto de Citações Normalizado da Categoria (NCI). Assim, áreas de pesquisa localizadas no quadrante superior direito onde a produção relativa e impacto de citações estão ambos acima de 1.0 são consideradas áreas mais fortes. Áreas de pesquisa localizadas no quadrante superior esquerdo onde a produção é relativamente baixa, mas o impacto de citações é alto são consideradas oportunidades potenciais. Uma área de pesquisa é considerada fraca se está localizada no quadrante inferior esquerdo onde tanto a produção relativa e o impacto de citações estão abaixo das médias mundiais.

Áreas de pesquisa no quadrante inferior direito com produções relativamente altas, mas com baixo impacto de citações estão sob ameaça.

Áreas de Forças e Oportunidades

Ciência da Saúde, com sua produção relativa de seis anos ligeiramente acima de 1,0 e o CNCI em 0,91 pode ser uma área potencialmente forte, especialmente porque os documentos publicados desde 2015 tiveram um impacto acima da média de citação no mundo. No canto inferior direito, áreas ameaçadas são Ciências Biológicas e Ciências Agrícolas, pois o Brasil realiza proporcionalmente mais pesquisas nessas áreas do que a média global. O Brasil é 3,5 vezes mais produtivo que a média mundial em Ciências Agrícolas, embora esses trabalhos tenham impacto de citação abaixo da média para o mundo e o Brasil. No entanto, a pesquisa agrícola, que geralmente é de tremenda importância local, nem sempre atrai a atenção internacional. Um CNCI baixo, como uma medida do impacto de citação global, não leva em consideração o impacto de pesquisas relevantes regionalmente.

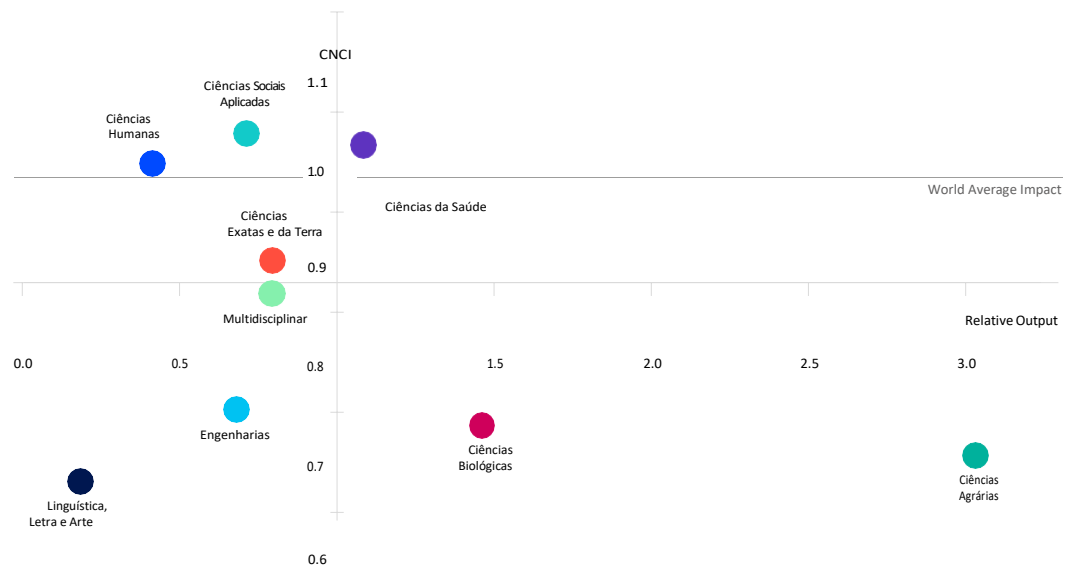
Linguística, Literatura e Artes e Engenharia no canto inferior esquerdo podem ser

consideradas áreas de pesquisa fracas. No entanto, como já observado, Linguística, Literatura e Artes, sua localização no canto inferior esquerdo pode ser um artefato de como a "contagem" é feita na *Web of Science*. Grande parte dessa pesquisa pode estar em português com um número limitado de leitores globais e, portanto, baixa cobertura na *Web of Science*. Assim, em termos globais, é provável que esta área apareça "Fracas" devido à sua natureza especializada e à pequena comunidade de pesquisa.

Com o Sirius18, a nova fonte de luz síncrotron brasileira, entrando em operação, existem mais oportunidades para pesquisas brasileiras nas Ciências Exatas e da Terra, onde o impacto da citação está acima da média para o Brasil, mas internacionalmente está proporcionalmente abaixo do que em outras áreas de pesquisa.

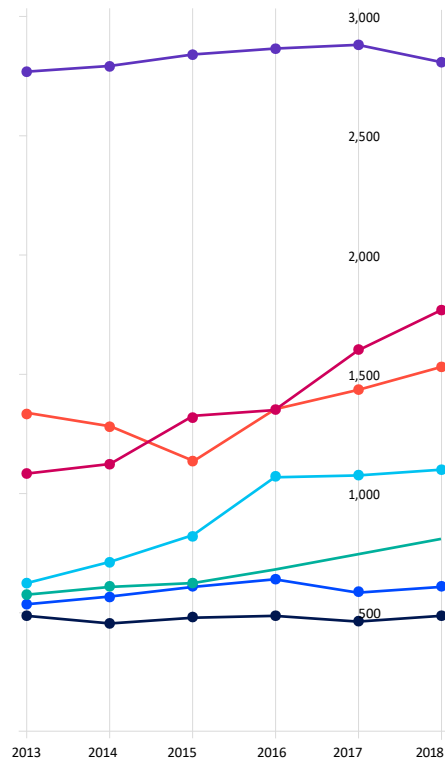
Por outro lado, análises adicionais devem ser conduzidas para verificar se as pontuações associadas às Ciências Sociais Aplicadas e Humanidades são artefatos de seu pequeno tamanho.

Figura 13



Forças, Fraquezas, Oportunidades e Ameaças na Pesquisa Brasileira de acordo com as áreas da CAPES.

Figure 14



Análises posteriores dos dados da subcategoria CAPES poderiam sugerir estratégias para aumentar o crescimento da pesquisa, citação de impacto ou alguma combinação de ambos. Por exemplo, a proporção de pesquisas em ciências agrárias no Brasil é três vezes superior à proporção global e continua a ser uma área altamente produtiva. Seu impacto na citação é persistentemente abaixo das médias mundiais, localizando-o em "ameaça" no canto inferior direito da Figura 13. No entanto, uma análise detalhada das subcategorias que compõem o campo poderia sugerir uma história diferente e estratégias diferentes para alcançar os objetivos de pesquisa do Brasil nesse campo.

O baixo CNCI das Ciências Agrícolas pode ser atribuído à Agronomia (CNCI de 0,46), responsável por mais de um terço de todos os trabalhos desta categoria de pesquisa (Figura 14). Mas o campo está mudando. Houve um rápido crescimento em áreas de pesquisa historicamente menos produtivas que foram alcançando ou que estão se aproximando do impacto médio da citação no mundo. O número de artigos cresceu 60% em Ciência e Tecnologia de Alimentos, que possui um CNCI de 1,11. Houve também um crescimento de artigos em Recursos de Pesca e Engenharia de Pesca cuja impacto de citação recentemente excedeu a média mundial. O nível de colaboração nas Ciências da Agricultura está baixo, em 21.7%, e abaixo da média do Brasil. Esforços para aumentar as colaborações internacionais poderiam aumentar a visibilidade internacional e potencialmente o CNCI.

Onde essa pesquisa é conduzida?

Quais são as organizações líderes em pesquisa no Brasil e onde elas se destacam?

O exame das cinco categorias de pesquisa da CAPES mais produtivas pode ser feito comparando a produção institucional e o impacto da citação de artigos publicados entre 2013 e 2018. A análise está dividida em duas categorias organizacionais: universidades e institutos de pesquisa especializados. São examinadas as 15 universidades e 10 institutos de pesquisa mais produtivos. Enquanto as universidades têm áreas com ênfase, a maioria realiza pesquisas em cada uma das cinco categorias de pesquisa da CAPES. Por outro lado, as instituições de pesquisa, com duas exceções notáveis, foram criadas para focar em áreas específicas de pesquisa, que são evidentes pela distribuição de suas pesquisas nas cinco categorias da CAPES.

A intensidade relativa da pesquisa em diferentes disciplinas é semelhante no Brasil e na maioria das universidades. Geralmente, as áreas mais produtivas de pesquisa para uma universidade são as áreas com maior produção em todo o Brasil; isto é, Ciências da Saúde, Ciências Biológicas e Ciências Exatas e da Terra. No entanto, parece haver universidades que fazem proporcionalmente mais pesquisas em ciências agrárias e em engenharia.

A Figura 15 mostra o número de trabalhos em cinco categorias da CAPES publicadas nessas 15 universidades brasileiras com a maior produção na *Web of Science* entre 2013 e 2018. O sombreamento indica o nível de produção da pesquisa, ou seja, tons mais escuros indicam instituições com maior produção dentro de uma categoria de pesquisa da CAPES. Essas 15 universidades produzem mais da metade de toda a produção de pesquisa no Brasil.

A Universidade de São Paulo domina entre as universidades, produzindo mais que o dobro do número total de publicações em comparação com a Universidade Estadual Paulista, que é a segunda universidade mais produtiva em termos de produção de pesquisa. Somente na área de Ciências Agrícolas

e, talvez em menor grau, Engenharia seja a distribuição da produção mais uniforme entre as universidades, onde outras universidades estão produzindo trabalhos em números comparáveis ao número de trabalhos publicados por pesquisadores afiliados à Universidade de São Paulo.

Figura 15

	Ciências da Saúde	Ciências Biológicas	Ciências Exatas e da Terra	Ciências Agrárias	Engenarias	
Universidade de Sao Paulo	21,912	17,025	14,536	6,476	6,819	58,899
Universidade Estadual Paulista	5,283	6,948	5,336	5,908	2,914	22,868
Universidade Estadual de Campinas	5,719	4,416	6,571	1,989	3,941	19,317
Universidade Federal do Rio de Janeiro	4,672	5,351	5,503	981	3,038	17,484
Universidade Federal do Rio Grande do Sul	5,199	4,009	3,960	2,168	2,599	15,860
Universidade Federal de Minas Gerais	5,233	4,349	3,293	1,809	2,108	14,904
Universidade Federal de Sao Paulo (UNIFESP)	7,372	3,186	1,212	358	724	11,228
Universidade Federal do Parana	2,133	3,333	2,486	2,190	1,628	9,995
Universidade Federal de Santa Catarina (UFSC)	2,473	1,974	2,468	1,358	2,284	9,162
Universidade Federal de Pernambuco	1,778	2,302	2,391	662	1,082	7,098
Universidade de Brasilia	1,756	2,039	2,023	895	892	7,056
Universidade do Estado do Rio de Janeiro	2,110	1,315	3,046	281	1,030	7,039
Universidade Federal de Sao Carlos	977	1,727	2,643	670	2,072	6,980
Universidade Federal de Viçosa	602	2,726	940	3,064	441	6,893
Universidade Federal de Santa Maria (UFSM)	1,247	1,809	1,425	2,522	782	6,670

Produção científica das 15 universidades mais produtivas em cinco categorias da CAPES e sua produção total nas 9 categorias.

Figura 16

	Ciências da Saúde	Ciências Biológicas	Ciências Exatas e da Terra	Ciências Agrárias	Engenharias	
Universidade Federal do ABC (UFABC)	1.06	0.74	1.95	-	0.95	1.68
Universidade Federal de São João del-Rei	0.65	0.61	2.53	1.29	0.95	1.54
Universidade Federal de Juiz de Fora	0.96	0.66	1.89	1.03	0.69	1.30
Universidade Federal de Sergipe	2.68	0.74	0.70	0.71	0.73	1.28
Universidade Federal de Pelotas	1.72	0.68	1.59	0.58	0.85	1.15
Universidade Federal de Santa Catarina (UFSC)	1.98	0.84	0.87	0.98	0.72	1.13
Universidade do Estado do Rio de Janeiro	0.91	0.72	1.37	0.63	1.09	1.06
Universidade Federal de São Paulo (UNIFESP)	1.17	0.94	0.73	0.68	0.87	1.06
Universidade Federal de Minas Gerais	1.56	0.90	0.80	0.93	0.79	1.03
Universidade Estadual de Campinas	0.86	0.94	1.23	0.87	0.74	1.03
Universidade de São Paulo	1.18	0.91	1.10	0.59	0.67	1.02
Universidade Federal do Rio Grande do Sul	1.50	0.78	0.97	0.88	0.82	1.02
Universidade Federal do Rio Grande do Norte	0.81	0.81	1.43	0.87	0.85	1.02
Universidade Federal do Rio de Janeiro	0.96	0.87	1.24	0.55	0.81	0.98
Universidade de Brasília	1.50	0.86	0.64	0.59	0.87	0.90

Impacto de citação normalizado da categoria das 15 universidades com maior impacto nas 5 categorias selecionadas e impacto global normalizados da pesquisa nas 9 categorias da CAPES.

A Figura 16 com impacto de citação mostra que as universidades mais produtivas nem sempre produzem as pesquisas mais citadas. Tons mais escuros na Figura 16 indicam um campo com maior CNCI e tons mais claros indicam menor CNCI naquela universidade. Uma universidade com um número insignificante de trabalhos em uma área é indicada por um "-". A Universidade Federal do ABC (UFABC) é a única nesta lista que possui insuficientes publicações em ciências agrícolas para calcular um CNCI.

Geralmente, as universidades têm uma ou duas áreas de alto impacto de citação que estão acima da média brasileira ou mundial. Com maior frequência, essas áreas são Ciências da Saúde e Ciências Exatas e da Terra,

que são os campos de maior impacto em todo o Brasil. As exceções estão acima do impacto médio mundial em Ciências Agrícolas da Universidade Federal de São João del-Rei e Engenharia na Universidade do Estado do Rio de Janeiro.

As universidades de pequeno e médio porte estão indo bem em termos de impacto de citação.

As universidades do estado de São Paulo estão presentes, mas não dominam essa lista de universidades classificada pelo CNCI geral, mostrado à direita na Figura 16. Como mencionado acima, a alta produção de pesquisa não está associada a um alto impacto de citação. As três principais universidades em produção de pesquisa (Figura 15) estão na metade inferior das universidades neste ranking de impacto de citação (Figura 16).

Figura 17

	Ciências da Saúde	Ciências Biológicas	Ciências Exatas e da Terra	Ciências Agrárias	Engenarias	
Empresa Brasileira de Pesquisa Agropecuária (Embrapa)	325	3,737	1,089	5,451	504	9,598
Fundação Oswaldo Cruz	5,264	4,983	663	560	138	9,195
Centro Brasileiro de Pesquisas Físicas	16	26	1,969	4	233	2,097
Instituto Nacional de Pesquisas da Amazônia	84	1,438	205	387	28	1,898
Instituto Nacional de Pesquisas Espaciais (INPE)	25	161	1,406	95	338	1,817
Comissão Nacional de Energia Nuclear (CNEN)	272	153	738	37	801	1,435
Hospital Israelita Albert Einstein	1,142	278	23	19	28	1,345
Instituto Butantan	450	1,046	90	99	31	1,294
Comando-Geral de Tecnologia Aeroespacial (CTA)	30	17	656	4	570	1,063
Instituto Tecnológico de Aeronáutica (ITA)	25	14	648	4	555	1,039

Produção de 10 Institutos de Pesquisa em cinco categorias da CAPES e Todas as Pesquisas em 9 categorias.

A Figura 17 mostra o número de artigos na *Web of Science* em cinco categorias CAPES, publicados por pesquisadores afiliados a esses 10 institutos de pesquisa entre 2013 e 2018. Apesar de sua especialização em uma ou no máximo duas áreas de pesquisa, a capacidade de pesquisa dos institutos de pesquisa é muito menor do que a capacidade das universidades. A produção da instituição de pesquisa mais produtiva, Empresa Brasileira de Pesquisa Agropecuária (Embrapa), é de menos de um sexto da produção de pesquisa da Universidade de São Paulo.

Como era de se esperar, a física domina a produção do Centro Brasileiro de Pesquisas Físicas e do Instituto

Nacional de Pesquisas Espaciais (INPE). Física e Engenharia são o foco da Comissão Nacional de Energia Nuclear (CNEN), do Comando Geral de Tecnologia Aeroespacial (CTA) e do Instituto Tecnológico de Aeronáutica (ITA) (Figura 17). Embora os trabalhos em Ciências Agrárias respondam por mais da metade da produção da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), também é um produtor substancial de pesquisas em Ciências Biológicas e, em menor grau, em Ciências da Terra e Exatas. Nenhum outro instituto de pesquisa parece ter um foco principal em Ciências Agrárias, exceto no Instituto Nacional de Pesquisas da Amazônia, no qual é a segunda maior categoria de pesquisa.

Figura 18

	Ciências da Saúde	Ciências Biológicas	Ciências Exatas e da Terra	Ciências Agrárias	Engenarias	
Centro Brasileiro de Pesquisas Físicas	-	-	2.23	-	0.80	2.21
Instituto Nacional de Pesquisas Espaciais (INPE)	-	1.49	2.52	0.54	0.95	2.19
Hospital Israelita Albert Einstein	1.75	1.09	-	-	-	1.62
Fundação Oswaldo Cruz	1.21	0.98	0.75	0.68	0.86	1.07
Instituto Nacional de Pesquisas da Amazônia	0.70	0.83	1.62	0.71	-	0.91
Empresa Brasileira de Pesquisa Agropecuária (Embrapa)	1.17	0.75	1.06	0.62	1.08	0.72
Instituto Butantan	0.80	0.70	0.58	0.66	-	0.68
Instituto Tecnológico de Aeronáutica (ITA)	-	-	0.78	-	0.60	0.65
Comando-Geral de Tecnologia Aeroespacial (CTA)	-	-	0.78	-	0.60	0.65
Comissão Nacional de Energia Nuclear (CNEN)	0.47	0.67	0.62	-	0.87	0.58

Categoria Impacto de Citação Normalizado (CNCI) das 10 principais instituições de pesquisa em cinco categorias da CAPES.

Na Figura 18, as instituições de pesquisa são classificadas por seu CNCI geral. Os dois institutos de pesquisa mais produtivos, o Centro Brasileiro de Pesquisas Físicas e o Instituto Nacional de Pesquisas Espaciais (INPE), concentram-se em Ciências Exatas e da Terra com o CNCI de 2,23 e 2,52, respectivamente. O CNCI geral de 2,21 e 2,19¹⁹ é significativamente maior do que o de qualquer universidade,

onde o CNCI mais alto é 1,68. No entanto, em termos de impacto de citação, as universidades têm maior profundidade, onde 13 das 15 universidades possuem CNCI acima da média mundial de 1,0, enquanto apenas quatro das dez instituições de pesquisa têm pontuação acima de 1,0. Assim como as universidades, os institutos de pesquisa com o maior número de publicações não apresentam os maiores impactos de citação.

¹⁹ It is likely that these high citation impacts are distorted because they include several massively multi-authored papers. For instance, there are 95 papers published with authors from INPE whose CNCI is 24.

Como o Brasil apoia a infraestrutura de pesquisa?

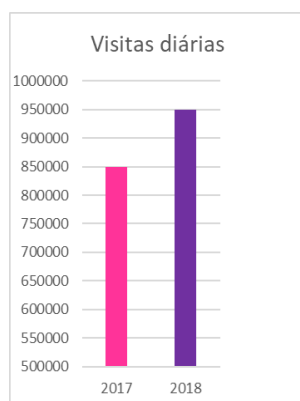
Como o Brasil aumenta a comunicação da pesquisa?

Além de contribuir para a expansão das artes, humanidades, engenharias e bases de conhecimento científico, além de treinar e educar a força de trabalho para a próxima geração, a criação de instituições e o suporte à infraestrutura científica também são contribuições valiosas para a empresa científica.

Em 1998, o Brasil lançou a Scientific Electronic Library Online (SciELO)²⁰, desenvolvida pela Fundação de Amparo a Pesquisa de São Paulo (FAPESP) em parceria com o BIREME, o Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde. Continua a florescer com o apoio adicional da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

A Coleção Scielo Brasil registra grande quantidade de consultas diárias. Em 2018 o número aproxima um milhão por dia (Figura 19²¹).

Figura 19



Acessos por dia a coleção Scielo Brasil 2017-2018

Com o multilinguismo²² como característica fundamental, a rede SciELO cobre e indexa a produção de pesquisas em português, espanhol e inglês da América Latina, Portugal, Espanha e África do Sul. De fato, 30% dos acessos registrados no Scielo Brasil em 2018 provêm do exterior²¹. Uma característica distintiva do SciELO é sua ênfase no acesso aberto e que todos os periódicos de sua coleção sejam indexados no Diretório de periódicos de acesso aberto (DOAJ). Os textos completos dos artigos do SciELO Brasil estão sendo baixados (download realizado) numa quantidade mensal que varia de 20 a 34 milhões, sendo a atividade acadêmica maior em junho e setembro e menor de dezembro a março (figura 20²¹)

Embora exista alguma sobreposição entre o SciELO e a coleção principal da *Web of Science*, o acesso ao SciELO Citation Index completo também pode ser obtido por meio das coleções regionais hospedadas na *Web of Science*.

Figura 20



Quantidade de downloads de texto inteiro por mês (Janeiro-Dezembro 2018)

²⁰ <https://scielo.org/>, <http://www.scielo.br/> accessed August 21, 2019

²¹ <https://analytics.scielo.org/w/accesses?collection=scl>, Packer http://discover.clarivate.com/Research_Excellence_Awards_Brazil_Download acessado 11/09/2019

²² Meneghini, R, Packer, A.L. (2007). Is there science beyond English? Initiatives to increase the quality and visibility of non-English publications might help to break down language barriers in scientific communication. *EMBO Reports*, 8(2): 112–116. Available at <http://bit.ly/2Mn4315>, accessed August 21, 2019

Descobertas e conclusões

Durante a janela de seis anos, 2013-2018, o volume da produção de pesquisa do Brasil continuou a crescer. O Brasil manteve sua posição como o 13º maior produtor mundial de publicações de pesquisa. Entre os países do BRICS, esse nível de produção está abaixo do da China e da Índia e acima do da Rússia e da África do Sul. Despesas domésticas brutas em pesquisa e Desenvolvimento (DRGE), como uma porcentagem do Produto Interno Bruto (PIB) do Brasil, oscilou entre 1,1 e 1,4%, de acordo com os dados mais recentes (2012- 2016) disponíveis²², que não atendem às expectativas de uma década atrás de que poderiam atingir dois por cento. Exceto pelos números potencialmente voláteis nos dois anos mais recentes, o percentual de artigos brasileiros entre o primeiro por cento dos artigos citados no mundo foi aumentando e se manteve próximo a um por cento.

A composição da pesquisa brasileira revela atividade e excelência concentradas em áreas que receberam investimentos direcionados ao setor. Por exemplo, embora as Ciências Agrárias como campo de pesquisa tenham baixo impacto da citação, desagregar isso em seus componentes constituintes mostra rápido crescimento em subcategorias que têm impacto acima da média mundial.

Investimentos adicionais nessas áreas aumentariam a visibilidade internacional e o impacto da citação. Da mesma forma, os investimentos em

projetos como o Sirius, a nova fonte de luz síncrotron brasileira, não apenas promovem a colaboração internacional, mas também podem atrair financiamento internacional para pesquisas.

As políticas que promovem colaborações universidade-indústria deram frutos, resultando em crescimento exponencial de trabalhos em coautoria com pesquisadores da indústria.

Um terço dessas colaborações são com empresas brasileiras e o restante com multinacionais. A Petrobras é a colaboradora mais forte das universidades e a indústria farmacêutica também tem uma longa história de colaboração.

As universidades públicas estão na vanguarda dessas colaborações com a indústria.

As organizações públicas também estão indo bem em termos de resultados de pesquisa e seu impacto na citação (CNCI). As 15 organizações mais produtivas - todas públicas - incluem 13 universidades e dois institutos de pesquisa especializados. Desses 15, 11 possuem um CNCI acima da média brasileira e oito têm um CNCI acima da média mundial.

Além de promover a pesquisa e a educação acadêmica, as políticas públicas brasileiras também apoiam a comunicação acadêmica, que rendeu à SciELO, a biblioteca eletrônica que cresceu além do Brasil para 13 outros países e em breve incluirá publicações acadêmicas de um total de 17 países.

Apêndice 1

Bibliometria e dados de citações

Bibliometria é sobre publicações e suas citações. O campo acadêmico surgiu da "ciência da informação" e agora geralmente se refere aos métodos usados para estudar e indexar textos e informações.

As publicações citam outras publicações. Esses links de citação crescem em redes e é provável que seus números estejam relacionados à importância ou ao impacto da publicação. O significado da publicação é determinado a partir de palavras-chave e conteúdo.

A análise de citações e a análise de conteúdo tornaram-se, portanto, uma parte comum da metodologia bibliométrica. Historicamente, métodos bibliométricos foram utilizados para traçar relações entre citações de periódicos acadêmicos. Agora, a bibliometria é importante na indexação do desempenho da pesquisa.

Os dados bibliométricos têm características particulares das quais o usuário deve estar ciente e são consideradas aqui.

Os artigos de periódicos (fontes de publicações) relatam trabalhos de pesquisa. Os artigos referem-se a

ou 'citam' o trabalho anterior relevante para o material que está sendo relatado. Novos artigos são citados por sua vez. Os artigos que acumulam mais citações são considerados como tendo maior "impacto", que é interpretado como significância ou influência em seu campo.

Portanto, a contagem de citações é reconhecida como uma medida de impacto, que pode ser usada para indexar a excelência da pesquisa a partir de um grupo, instituição ou país específico.

As origens da análise de citação como uma ferramenta que poderia ser aplicada ao desempenho da pesquisa podem ser encontradas em meados da década de 1950, quando Eugene Garfield propôs o conceito da indexação de citações e introduziu o Science Citation Index, o Social Sciences Citation Index e o Arts & Humanities Citation Index, produzidos pelo Institute of Scientific Information (agora parte da Clarivate Analytics).

Podemos contar citações, mas elas são apenas "indicadores" de impacto ou qualidade - não métricas. A maioria dos indicadores de impacto usa contagens médias de citações de grupos de artigos, porque alguns documentos individuais podem ter perfis de citações incomuns ou enganosos. Esses valores extremos são diluídos em amostras maiores.

Fonte de dados

Os dados utilizados nesta análise vieram do banco de dados da *Web of Science*, que dá acesso não apenas a periódicos, mas também a anais de conferências, livros, patentes, sites e estruturas químicas, compostos e reações. A *Web of Science* possui uma estrutura unificada que integra todos os dados e termos de pesquisa juntos e, portanto, fornece um nível de comparabilidade não encontrado em outros bancos de dados. É amplamente reconhecida como a principal fonte mundial de citações e dados bibliométricos. A *Web of Science* concentra-se em pesquisas publicadas em periódicos, conferências e livros de ciências, medicina, artes, humanidades e ciências sociais.

A *Web of Science* foi criada originalmente como uma ferramenta de conscientização e recuperação de informações, mas adquiriu um importante uso primário como ferramenta de avaliação de pesquisas, usando análise de citações e bibliometria. A cobertura de dados é atual e retrospectiva nas ciências, ciências sociais, artes e humanidades, em alguns casos desde 1900. Na comunidade de pesquisa, essa fonte de dados era anteriormente referida pela sigla 'ISI'.

Ao contrário de outros bancos de dados, a *Web of Science* e os bancos de dados subjacentes são seletivos, ou seja: os periódicos indexados são selecionados usando rigorosos critérios editoriais e de qualidade. O conteúdo qualificado e multidisciplinar cobre mais de 18.000 periódicos de maior impacto em todo o mundo, incluindo periódicos de acesso aberto e mais de 180.000 anais de conferências. Os periódicos indexados abrangem a maioria dos relatórios científicos significativos citados com frequência e, mais importante, uma proporção ainda maior da produção científica citada. Esse processo seletivo garante que a contagem de citação permaneça relativamente estável em determinados campos de pesquisa e não oscile indevidamente de ano para ano, o que aumenta a usabilidade desses dados para avaliação de desempenho.

A *Web of Science Group* possui uma vasta experiência com bancos de dados sobre entradas, atividades e produtos de pesquisa e desenvolveu abordagens analíticas inovadoras para comparar e interpretar o impacto de pesquisas internacionais, nacionais e institucionais.

22

Existem 22 campos no *Essential Science Indicators* e 254 campos na *Web of Science*.

Categorias nas bases de dados

Os dados de origem podem ser agrupados em vários sistemas de classificação. A maioria deles é baseada em grupos de periódicos que têm uma ligação de citação cruzada relativamente alta e se agrupam naturalmente. Classificações especiais usam mapas de assuntos em dados de terceiros, como as categorias da OCDE definidas no manual Frascati.

O Grupo *Web of Science* freqüentemente usa as categorias de campo mais amplas no sistema *Essential Science Indicators* e as categorias mais refinadas de periódicos na *Web of Science*. Existem 22 campos em *Essential Science Indicators* e 254 campos na *Web of Science*. Em ambos os casos, nossas análises bibliométricas se baseiam em toda a gama de metadados disponíveis no banco de dados subjacente; portanto, as análises em nossos relatórios diferem um pouco de qualquer análise criada instantaneamente na interface em linha.

A maioria das análises começa com uma visão geral dos dados, depois passa para uma visão de categorias amplas e só depois se concentra em um nível mais refinado nas áreas de maior interesse para políticas, programas ou propósitos institucionais.

Atribuindo artigos a instituições

Um artigo é atribuído a cada país e a cada instituição cujo endereço aparece pelo menos uma vez para qualquer autor desse artigo. Um artigo conta uma vez e apenas uma vez para cada tarefa, no entanto, muitas variantes de endereço podem ocorrer para o país ou instituição. Nenhuma ponderação é aplicada.

Por exemplo, um artigo tem cinco autores, assim:

Author	Institution	Country		
Gurney, KA	Univ Leeds	UK	Counts for Univ Leeds	Counts for UK
Adams, J	Univ Leeds	UK	No gain for Univ Leeds	No gain for UK
Kochalko, D	Univ C San Diego	USA	Counts for UCSD	Counts for USA
Munshi, S	Gujarat Univ	India	Counts for Gujarat Univ	Counts for India
Pendlebury, D	Univ Oregon	USA	Counts for Univ Oregon	No gain for USA

Portanto, este artigo com cinco autores seria incluído uma vez na conta de cada uma das quatro universidades e uma vez na conta de cada um dos três países.

O trabalho realizado no Grupo *Web of Science* e pesquisas publicadas em outros lugares indicam que a ponderação fracionada com base no saldo dos autores por instituição e país faz pouca diferença para as conclusões de uma análise

em nível agregado. Essa análise fracionária pode introduzir erros imprevistos na tentativa de criar uma tarefa detalhada, mas incerta.

O particionamento de crédito faria uma diferença maior em um nível detalhado de grupos de pesquisa, mas a análise pode então ser validada manualmente.

Contagem de citações

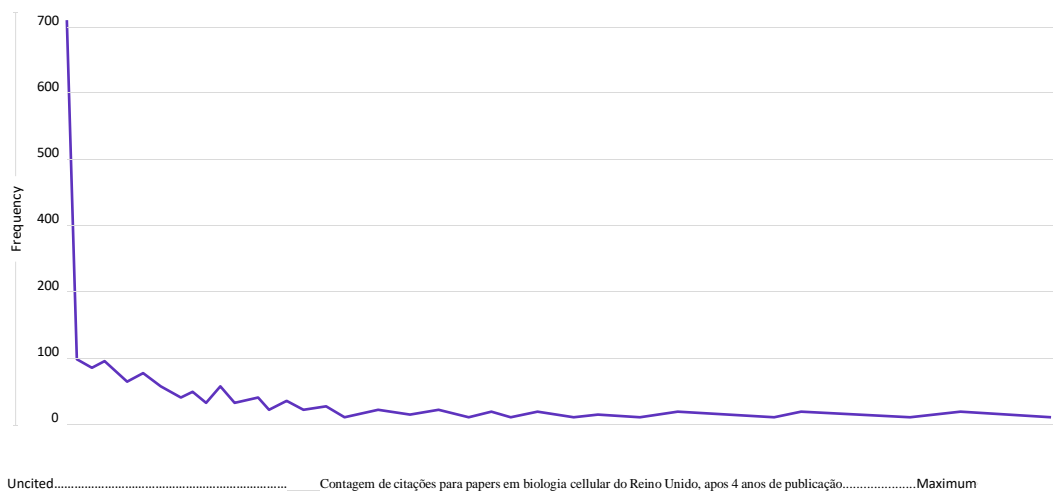
Uma publicação acumula contagens de citações quando é referida por publicações mais recentes. Alguns trabalhos são citados com frequência e muitos são citados raramente ou nunca, portanto a distribuição de citações é altamente distorcida.

Por quê há tantos artigos não citados?

Certamente, alguns trabalhos permanecem não citados porque seu conteúdo tem pouco ou nenhum impacto, mas esse não é o único motivo. Pode ser porque eles foram publicados em um periódico não lido por pesquisadores para quem o artigo pode ser interessante. Pode ser que eles representem relatórios de trabalho importantes, mas "negativos"; um beco sem saída a ser evitado por outros. A publicação pode ser um comentário em um editorial e não um artigo de revista normal e, portanto, de interesse geral e não de pesquisa, ou pode ser que o trabalho seja uma 'bela adormecida' que ainda pode ser reconhecido por seu significado.

Outros artigos podem ser muito citados: centenas, até milhares de vezes. Novamente, existem várias razões para isso. O trabalho mais frequentemente citado está sendo reconhecido por seu significado e impacto inovadores no campo de pesquisa de que fala. O impacto aqui é um bom reflexo da qualidade: é um indicador de excelência. Mas existem outros trabalhos que são frequentemente citados porque seu significado é um pouco diferente: eles descrevem a metodologia-chave; eles são uma revisão cuidadosa e abrangente de um campo; ou representam visões contenciosas que outros procuram refutar.

A análise de citações não pode fazer julgamentos de valor sobre porque um artigo é citado nem porque ele é altamente citado. A análise pode apenas relatar o impacto da citação que a publicação alcançou. Normalmente assumimos, com base em muitos outros estudos que vinculam julgamentos bibliométricos e de pares, que alta contagem de citações se correlaciona, em média, com a qualidade da pesquisa.



A figura mostra a distribuição distorcida dos artigos citados com mais ou menos frequência de uma amostra de publicações de autoria do Reino Unido em biologia celular. A inclinação na distribuição varia de campo para campo. É para compensar esses fatores que as contagens reais de citações devem ser normalizadas ou redimensionadas contra uma linha de base mundial.

Não procuramos explicar separadamente o efeito da autocitação. Se a contagem de citações for significativamente afetada pela autocitação, é provável que o artigo não tenha sido citado com muita frequência. Portanto, isso é apenas uma consequência para atividades de baixo impacto. Estudos mostram que para grandes amostras em nível nacional e institucional o efeito da autocitação tem pouco ou nenhum efeito sobre os resultados analíticos e não alteraria a interpretação dos resultados.

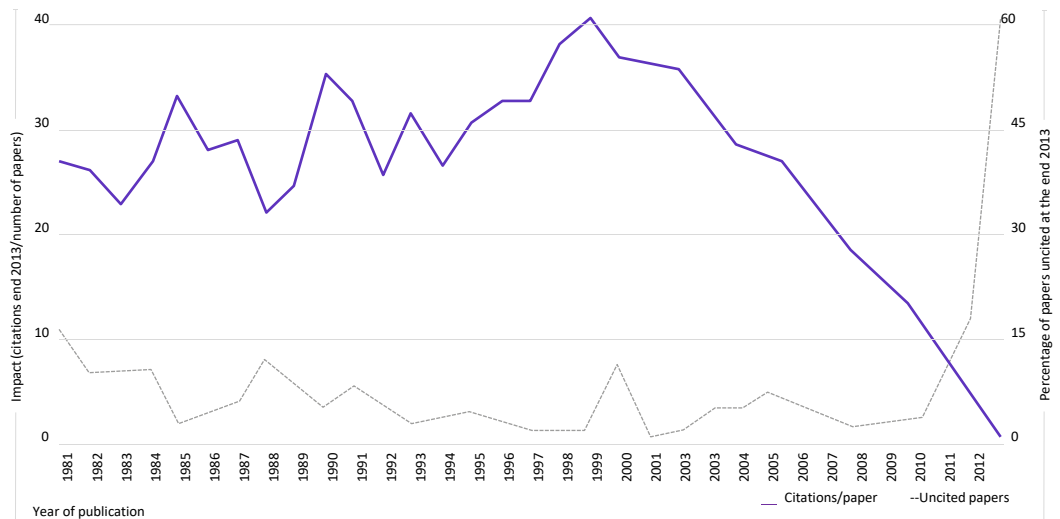
Fatores Temporais

As citações se acumulam com o tempo. Os artigos mais antigos, portanto, têm, em média, mais citações do que trabalhos mais recentes. O gráfico abaixo mostra o padrão de acumulação de citações para um conjunto de 33 revistas na categoria de periódicos Ciência dos Materiais, Biomateriais.

Trabalhos com menos de oito anos ainda estão, em média, acumulando citações adicionais. A contagem de citações alcança um platô para fontes mais antigas.

O gráfico mostra que a porcentagem de artigos que nunca foram citados cai em cerca de cinco anos. Além de cinco anos, entre 5% e 10% dos papéis permanecem não citados.

É preciso levar em consideração esses fatores de tempo na comparação da pesquisa atual com os padrões históricos. Por esses motivos, às vezes é mais apropriado usar uma janela fixa de cinco anos de artigos e citações para comparar dois períodos do que observar o perfil de longo prazo das citações e a falta de citações para um ano recente e um ano histórico.



Fatores disciplinares

As taxas de citação variam entre disciplinas e campos. Para a base científica do Reino Unido como um todo, dez anos produzem um platô geral além do qual seriam esperadas poucas citações adicionais. No conjunto, as citações se acumulam mais rapidamente e atingem um nível mais elevado nas ciências biológicas do que nas ciências físicas, e as ciências naturais geralmente tem taxas de citações mais altas que as ciências sociais.

Os artigos são atribuídos a disciplinas (categorias de periódicos ou campos de pesquisa) pela *Web of Science*, reunindo áreas de pesquisa cognatas. Antes de 2007, os periódicos eram atribuídos às categorias mais antigas e bem estabelecidas do *Current Contents*, que foram informadas por um extenso trabalho de Thomson e com a comunidade de pesquisa desde o início dos anos 1960. Esse esquema foi substituído pelas 254 categorias de periódicos da *Web of Science*, que permitem maior desagregação para o crescente volume de pesquisa publicada e indexada.

Os artigos são alocados de acordo com a revista em que o artigo é publicado.

Alguns periódicos podem ser considerados parte do registro de publicação em mais de um campo de pesquisa.

Como o exemplo abaixo ilustra, o periódico *Acta Biomaterialia* é atribuído a duas categorias de periódicos: Ciência dos Materiais, Biomateriais e Engenharia Biomédica.

Muito poucos artigos não são atribuídos a nenhum campo de pesquisa e, como tal, não serão incluídos em análises específicas usando dados de impacto de citação normalizados. Os periódicos incluídos nos bancos de dados da *Web of Science* e como são selecionados são detalhados aqui: clarivate.com/webofsciencelgroup/solutions/webofscience-core-collection-editorial-selection-process/

Alguns periódicos com um conteúdo muito diversificado, incluindo os prestigiados periódicos *Nature* e *Science*, foram classificados como Multidisciplinares em bancos de dados criados antes de 2007. Os trabalhos desses periódicos multidisciplinares agora são redistribuídos para campos de pesquisa mais específicos usando um algoritmo baseado na(s) área(s) de pesquisa das referências citadas no artigo.

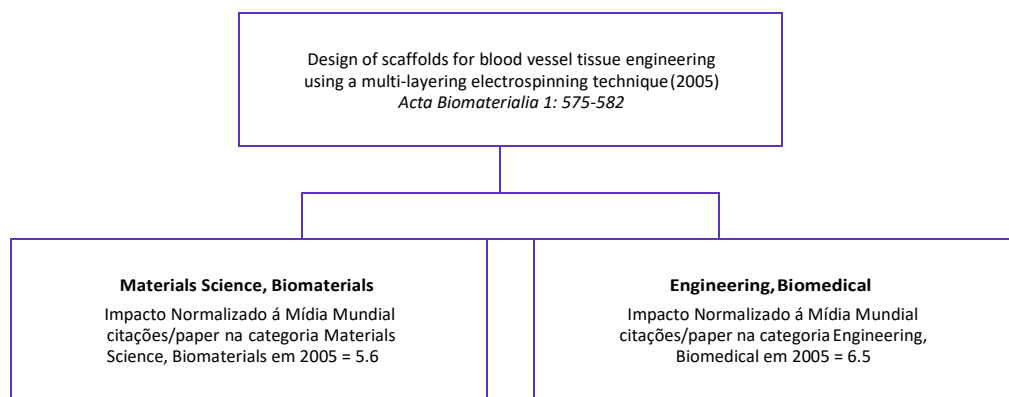
Impacto de citação normalizado

Como as citações se acumulam ao longo do tempo a uma taxa dependente do campo de pesquisa, todas as análises devem levar em consideração o campo e o ano. Em outras palavras, porque a contagem absoluta de citações para um artigo específico é influenciado por seu campo e pelo ano em que foi publicado, só podemos fazer comparações de dados indexados após normalizar com referência a essas duas variáveis.

Utilizamos apenas contagens de citações para revisões e artigos nos cálculos de impacto, porque o tipo de documento influencia a contagem de citações. Por exemplo, uma revisão geralmente será

citada mais frequentemente do que um artigo no mesmo campo, mas editoriais e resumos de reuniões raramente são citados e as taxas de citação para procedimentos de conferência são extremamente variáveis. Os fatores de normalização mais comuns são as citações médias por artigo para (1) o ano e (2) o campo ou a revista em que o artigo foi publicado. Essa normalização também é referida como "rebasear" a contagem de citações.

Portanto, o impacto é mais comumente analisado em termos de 'impacto de citação normalizado' ou CNCI. O esquema a seguir ilustra como o impacto de citação normalizado é calculado no nível do artigo e da(s) categoria(s) do periódico.



Este artigo da revista *Acta Biomaterialia* é atribuído a duas categorias de revistas: Ciência dos Materiais, Biomateriais e Engenharia Biomédica. As linhas de base da média mundial para, por exemplo, Ciência dos materiais, Biomateriais são calculadas somando as citações a todos os artigos e revisões publicadas mundialmente na revista *Acta Biomaterialia* e as outras 32 revistas atribuídas a esta categoria para cada ano, dividindo-a pelo número total de artigos e revisões publicados na categoria de periódicos. Isso fornece o impacto da citação normalizada específica da categoria (no exemplo acima, o CNCI específico da categoria para Ciência dos Materiais, Biomateriais é 5,6 e o CNCI específico da categoria para Engenharia Biomédica é maior em 6,5).

A maioria dos trabalhos (quase dois terços) é atribuída a uma única categoria de periódico, enquanto uma minoria deles é atribuída a mais de cinco.

O impacto médio (normalizado) da citação pode ser calculado em um nível de artigo individual, onde pode ser associado a mais de uma categoria de periódico. Também pode ser calculado para um conjunto de artigos em qualquer nível, de um único país até a produção de um pesquisador individual. No exemplo acima, o impacto médio da citação do artigo da *Acta Biomaterialia* pode ser expresso como $((5,6 + 6,5) / 2) = 6,1$.

Os dados de impacto médio mundial são provenientes dos dados da linha de base dos Indicadores Nacionais de Ciências da *Web of Science* para 2015.

Impacto de citação médio normalizado

O desempenho da pesquisa tem sido historicamente indexado usando o impacto médio da citação, geralmente comparado à média mundial que toma conta de tempo e disciplina. Como observado, no entanto, a distribuição de citações entre artigos é altamente distorcida, porque muitos artigos nunca são citados, enquanto alguns artigos acumulam contagens de citações muito grandes. Isso significa que uma média pode ser enganosa se forem feitas suposições sobre a distribuição dos dados subjacentes.

De fato, quase todas as métricas de atividade de pesquisa são distorcidas: para renda de pesquisa, números de PhD e publicações, existem muitos valores baixos de atividade e alguns valores excepcionalmente altos. Na realidade, portanto, a distribuição distorcida significa que o impacto médio tende a ser maior e muitas vezes significativamente diferente da mediana ou do modo na distribuição. Isso deve ser lembrado ao revisar os resultados analíticos.

Perfis de impacto

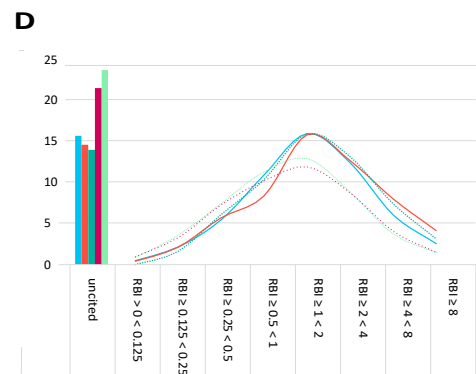
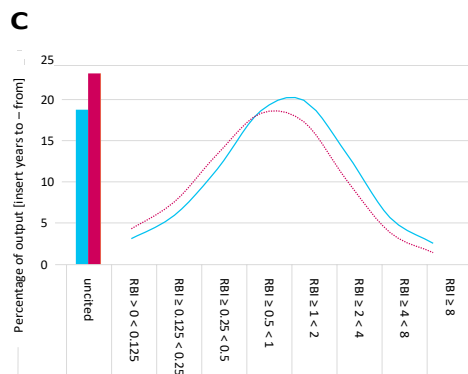
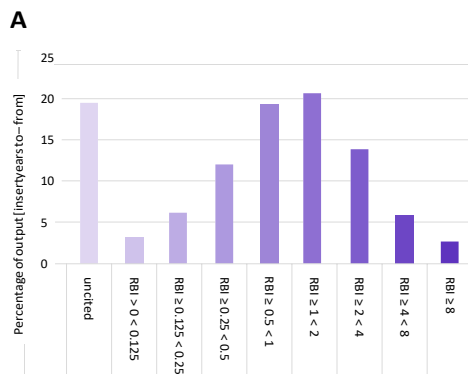
Desenvolvemos uma metodologia bibliométrica que mostra a proporção de artigos não citados e a proporção que se encontra em cada um dos oito ranges

das taxas de citação relativas, normalizadas (rebaseadas) para a média mundial²³. Um Perfil de Impacto permite um exame e análise dos pontos fortes e fracos das produções publicadas em relação à média mundial e em relação a um perfil de referência. Isso fornece muito mais informações sobre a base e a estrutura do desempenho da pesquisa do que as médias relatadas convencionalmente nos índices de citação.

Os artigos que são "altamente citados" são frequentemente definidos em nossos relatórios como aqueles com um impacto médio de citação maior que ou igual a 4,0, ou seja, aqueles artigos que receberam maior ou igual a quatro vezes o número médio mundial de citações para artigos nesse assunto publicado naquele ano. Isso difere do banco de dados de Artigos Altamente Citados, produzidos pela *Web of Science*, que são os um por cento mais frequentemente citados em seu campo e ano. O percentil superior é um poderoso indicador de desempenho líder, mas é um limite rigoroso demais para a maioria das análises de desempenho.

A proporção de artigos não citados em um conjunto de dados pode ser comparada à referência para o Reino Unido, os EUA ou qualquer outro país. No geral, em uma amostra típica de dez anos, cerca de um quarto dos artigos não foram citados no período de 10 anos; a maioria deles é, obviamente, de artigos que foram publicados mais recentemente.

O histograma do perfil de impacto pode ser apresentado de várias maneiras ilustradas abaixo:



A: é usado para representar a produção total de um país, instituição ou pesquisador, sem dados de referência. Visualmente, destaca o número de artigos não citados (pontos fracos) e artigos altamente citados (pontos fortes).

B & C: são usados para representar a produção total de um país, instituição ou pesquisador (cliente) em um conjunto de dados de benchmark apropriado (benchmark). Os dados são exibidos como histogramas (B) ou uma combinação de histograma e perfil

C: a versão C evita a 'viagem' que ocorre em histogramas em que

o olho é atraído para os dados mais deslocados para a direita, mas pode ser mais difícil de interpretar como dados categóricos.

D: ilustra a complexidade dos dados que podem ser exibidos usando um perfil de impacto. Esses dados mostram os resultados da pesquisa em categorias definidas de periódicos com relação aos parâmetros de referência adequados:

- cliente, campo de pesquisa X (azul claro);
- cliente, campo de pesquisa Y (laranja);
- cliente, campo de pesquisa Z (verde);
- referência, campo de pesquisa X + Y (roxo);
- referência, campo de pesquisa, Z (verde claro).

Os Perfis de Impacto permitem um exame e análise do balanço da produção publicada em relação à média mundial e em relação a um perfil de referência. Isso fornece muito mais informações sobre a base e a estrutura do desempenho da pesquisa do que as médias relatadas convencionalmente nos índices de citação.

Um Perfil de Impacto mostra qual a proporção de artigos não citados e qual a proporção em cada uma dos oito ranges de taxas de citação relativas, normalizadas para a média mundial (que se torna 1,0 neste gráfico). Taxas de citação normalizadas acima de 1,0 indicam trabalhos citados com mais frequência do que a média mundial para o campo em que a revista é categorizada e em seu ano de publicação.

Deve-se prestar atenção a:

- A proporção de artigos não citados à esquerda do gráfico
- A proporção de artigos citados em • A localização do grupo (modal) mais comum perto do centro
- A proporção de artigos nas categorias mais citadas à direita (≥ 4 x mundo, ≥ 8 x mundo). ambos os lados da média mundial (1,0)

O que são artigos não citados?

Pode ser uma surpresa que alguns artigos de periódicos nunca sejam citados posteriormente após a publicação, mesmo por seus autores. Isso representa cerca de metade da produção global total de um período típico recente de 10 anos

Qual é o limite para "altamente citado"?

O Grupo Web of Science tradicionalmente usa o termo "Artigo altamente citado" para se referir ao 1% do mundo dos artigos mais citados, considerando o ano de publicação e o campo. Após analisar os resultados de várias análises, escolhemos uma definição mais relaxada para o nosso trabalho descritivo e analítico. Consideramos papéis que estão entre os 10% melhores dos artigos mais citados no mundo, considerando o ano de publicação e o campo, sendo relativamente altamente citados para comparações nacionais.

Apêndice 2

Metodologia para a colaboração universidade-indústria

A análise aqui apresentada utiliza dados da *Web of Science*, obtidos através de pesquisas realizadas na interface WoS disponível para pesquisadores.

Enquanto o banco de dados InCites carrega dados para a porcentagem de artigos com coautoria do setor, seus dados estão incompletos, pois o banco de dados ainda não é capaz de classificar corretamente a natureza de um grande número de organizações empresariais no Brasil (e em outros lugares, nesse caso). Para obter os dados mostrados aqui, desenvolvemos uma rotina de busca especialmente criada para desvendar o setor de negócios no Brasil.

O procedimento envolveu a obtenção dos dados para todos os documentos científicos no banco de dados com pelo menos um autor no Brasil

(> 300.000 registros), analisando as organizações às quais os autores estavam afiliados (> 22.000) e, em seguida, classificando entre as que estavam no setor de negócios. No final, tínhamos mais de quatro mil organizações. Nesse ponto, fizemos uma pesquisa procurando itens nos quais os autores estavam em uma das mais de 4.000 organizações do setor empresarial e em toda e qualquer universidade (obtida em uma lista separada).

Obter dados completos é um desafio. A *Web of Science* (e Scopus) possui uma classificação incompleta das organizações do setor comercial, de modo que seu indicador sobre as coautorias entre universidades e indústria subestimam o tamanho real das colaborações. O banco de dados atualmente acessível é categorizado como “indústria”, principalmente empresas multinacionais e possui fraca (ou nenhuma) categorização de médias e pequenas empresas brasileiras²⁴.

²⁴ Esse fato não deve ser interpretado como uma crítica às bases bibliométricas. Se alguma coisa, podemos ter críticas sobre a maneira como os bancos de dados são usados com frequência. Esses bancos de dados foram desenvolvidos inicialmente para ajudar a comunidade científica a encontrar referências úteis para usar em seus trabalhos. Eles são herdeiros da tradição iniciada pelo Science Citation Index, criado por Eugene Garfield, como um banco de dados de informações sobre publicações, classificadas por setor, autor, tópico. Muitos de nós, cientistas, lembramos das viagens à biblioteca para consultar o Science Citation Index (e o Current Abstracts) para tentar acompanhar as novas publicações em nossos campos. Parte do desafio aqui é que os bancos de dados não foram construídos para serem instrumentos de classificação. Ainda assim, eles podem auxiliar na classificação ou estudos de políticas científicas, se usados com o devido cuidado.

A *Web of Science* oferece, no entanto, ferramentas de pesquisa que podem ser usadas para compensar pela falta de categorização. Utilizamos a seguinte metodologia:

1. Começamos com uma busca por todos os itens com endereços no “Brazil” (ou “Brasil”), obtendo uma lista de aproximadamente 700.000 itens.

2. Em seguida, usamos as “Ferramentas de Análise” para obter a lista de instituições mencionadas no campo “Endereço” de cada publicação. Essa lista tinha 22.000 entidades.

3. A lista de entidades foi analisada linha a linha para identificar aquelas pertencentes ao setor de negócios que foram usadas para compor uma sequência de pesquisa aplicada para o campo "Nome da organização".

4. Além da lista de termos de pesquisa obtidos em (3), adicionamos à cadeia de itens os itens que terminam em terminações que caracterizam entidades do setor comercial, como LTDA, LLC, SpA, LTD, BV, INC e outras, usando caracteres curinga para a parte inicial do nome da organização.

5. Adicionamos à lista resultante de termos de pesquisa os nomes de organizações no Brasil conhecidas por terem atividades de P&D

(por exemplo, empresas financiadas pelo Programa de Pesquisa Inovadora para Pequenas Empresas da FAPESP, empresas listadas na lista do jornal Valor Econômico das empresas mais inovadoras do Brasil). Isso nos permitiu adicionar mais de 200 nomes de organizações à sequência de pesquisa.

Finalmente, adicionamos os nomes de todas as entidades consideradas pelo InCites como empresas.

A lista de entidades tinha mais de 4.000 itens neste momento.

Em seguida, foi necessário eliminar redundâncias para cumprir com a restrição da ferramenta de pesquisa da *Web of Science* que limita o número de termos de pesquisa a cerca de 6.000, aparentemente considerando os operadores booleanos no cálculo. Muitas empresas aparecem sob nomes semelhantes. Em outros casos, era necessário consolidar nomes: por exemplo, os trabalhos da Petrobras CENPES aparecem sob vários nomes de entidades como 'Leopoldo Miguez', Cenpes Petrobras, 'Centro Leopoldo M' e outras combinações.

Em seguida, a sequência de pesquisa foi criada com os nomes das empresas e os nomes das universidades no Brasil.

Sobre o Grupo *Web of Science*

O Grupo *Web of Science* organiza as informações de pesquisa do mundo para permitir que academias, corporações, editores e governos acelerem o ritmo da pesquisa. Ele é alimentado pela *Web of Science*, plataforma de inteligência de pesquisa e o maior índice mundial de citações independente de editoras. Suas muitas marcas conhecidas também incluem *Converis*, *EndNote*, *Kopernio*, *Publons*, *ScholarOne* e o *Institute for Scientific Information (ISI)*. A "universidade" do Grupo *Web of Science*, ISI mantém o corpus de conhecimento sobre o qual são construídos o índice e informações relacionadas, conteúdo e serviços analíticos; divulga esse conhecimento externamente através de eventos, conferências e publicações e realiza pesquisas para sustentar, ampliar e melhorar a base de conhecimento. O Grupo *Web of Science* é uma empresa da *Clarivate Analytics*.

Contate nossos especialistas hoje:

+1 215 386 0100 (E.U.A)

+44 (0) 20 7433 4000 (Europa)

+ 55 (11) 94302 5492 (Brasil)

webofsciencgroup.com

© 2019 Clarivate Analytics. Web of Science Group and its logo, as well as all other trademarks used herein are trademarks of their respective owners and used under license.