

DADOS ABERTOS DO RANKING DE LEIDEN: DIVERSIDADE DE ECOSISTEMAS DE UNIVERSIDADES

Peter A. Schulz

Professor Titular

Universidade Estadual de Campinas (Unicamp)

E-mail: Peter.schulz@fca.unicamp.br

INTRODUÇÃO

Indicadores bibliométricos são componentes importantes na composição de *rankings* de universidades, notadamente os números de artigos publicados em uma determinada base de dados, como a Web of Science ou a Scopus; bem como medidas de impacto dessa produção científica, derivadas diretamente dos números de citações recebidas pelos artigos indexados. Esses indicadores são integrados a outros, com diferentes pesos, compondo um indicador final que ordenará as universidades em uma lista, que são como os *rankings* globais de universidades são apreciados por um público crescente, desde seu surgimento. Além desses indicadores de produção e impacto, outros são possíveis de serem minerados dos metadados das bases mencionadas. Em especial, para o interesse do presente trabalho, indicadores de colaboração e acesso aberto começaram a ser disponibilizados.

No entanto, a maioria dos *rankings* não publicita os dados utilizados para sua composição, tornando difícil a avaliação de suas metodologias. *Rankings* de universidades são “dispositivos calculistas”, ou seja, “tornam entidades altamente diversas – como as universidades – mensuráveis por meio de métodos quantitativos, ou, dito de outra forma, *rankings* tornam as universidades ‘calculáveis’”. (HAMMARFELT; RIJKE; WOUTERS, 2017) Essa calculabilidade é um processo com três etapas. Primeiro, as universidades são desconectadas de seu contexto original, com sua complexidade transformada em números, transferíveis e comunicáveis para um público mais amplo e permitindo novas contextualizações. Em um segundo passo, esses elementos descontextualizados (os indicadores) são comparados e articulados entre si. Em uma terceira etapa, uma lista precisa ser produzida, através de uma integração dos indicadores, e apresentada na forma de um *ranking* uniforme (mas de entidades dessemelhantes).

Nesse contexto, o objetivo deste trabalho é identificar e analisar as dessemelhanças entre instituições e a existência de diferentes ecossistemas de universidades (LEYDES-

DORFF; BORNMANN; MINGERS, 2019) colocadas em listas (os *rankings*) que escondem essa heterogeneidade. Esse objetivo é possível a partir dos dados abertos que compõem o *Ranking* de Leiden de 2019, gerado pelo Centro de Estudos em Ciência e Tecnologia (CWTS) da Universidade de Leiden, Holanda. Esse *ranking* apresenta uma particularidade importante: considera exclusivamente a pesquisa a partir de artigos publicados, que constam na base Web of Science, ou seja, refere-se a um recorte de apenas uma das missões da universidade. A edição de 2019 contempla 963 universidades que tenham publicado pelo menos 1000 artigos, somando as diferentes áreas do conhecimento no período 2014-2017. Esse recorte, é importante lembrar, não permite discutir a relação entre a pesquisa e as outras missões das universidades – ensino e extensão – elementos definidores do papel das universidades em diferentes sistemas regionais e nacionais. No entanto, os resultados apresentados evidenciam que, mesmo dentro do recorte analisado, as diferenças entre os sistemas são significativas sugerem pistas para diferentes usos da pesquisa científica em diferentes países. O *Ranking* de Leiden é disponibilizado em um portal, que não apresenta listas derivadas de integração de indicadores, mas sim a partir de cada indicador, separados em quatro categorias: impacto, colaboração, acesso aberto e gênero (essa última não será discutida nesse trabalho). Apesar dessa diferença na apresentação dos dados, o *ranking* de Leiden é percebido como os outros *rankings*,¹ tanto por acadêmicos, quanto pelo público geral. Esse equívoco de apreciação assemelha-se muito ao que ocorre com os próprios indicadores bibliométricos pela chamada “cientometria leiga”. (LEYDESDORFF; WOUTERS; BORNMANN, 2016) A característica mais importante, no entanto, é a disponibilização do acervo de dados que podem ser descarregados livremente, possibilitando diversos cruzamentos e verificações de hipóteses e correlações, que constituem o corpo central do presente trabalho.

METODOLOGIA & DADOS ABERTOS

A abordagem considerada nesse trabalho é a de empregar uma metodologia aberta, ou seja, baseada no uso de dados abertos, permitindo a verificação dos resultados. Além disso, visa proporcionar, de maneira mais direta possível, a adoção dos protocolos utilizados para seleções de outros grupos de universidades, áreas do conhecimento e indicadores pelos leitores interessados. Complementando essa perspectiva, todos os tratamentos de dados são por meio das ferramentas disponíveis na planilha Excel. A planilha disponibilizada livremente pelo *Ranking* de Leiden 2019 fornece mais de 80 indicadores para as 963 universidades, que podem ser filtradas por país (56 países) e áreas do conhecimento (Ciências Biomédicas e da Saúde, Ciências da Terra e da Vida, Matemática e Ciências da Computação, Ciências Físicas e engenharia e Ciências Sociais e Humanidades).

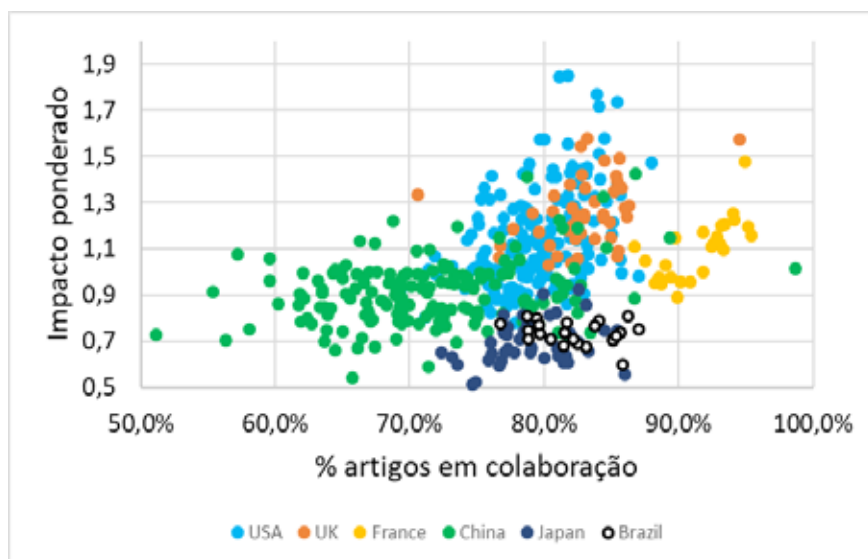
¹ Ver: <https://www.leidenranking.com/downloads>.

Os resultados apresentados aqui são baseados em indicadores independentes do tamanho das universidades e considerando contagem fracionária, isto é, ponderados pelo número de universidades colaborantes em cada artigo. Como indicador de impacto foi escolhido o MNCS: média de citações ponderada por área de conhecimento e ano de publicação. Para indicador de colaboração foi escolhido o PP(collab), a proporção das publicações de um universidade em coautoria com outras organizações, e o PP(collab int), que é o indicador equivalente apenas para colaborações internacionais. O indicador de acesso aberto considerado é o PP(AO): proporção de documentos publicados de cada universidade com algum tipo de acesso aberto. Estes indicadores foram selecionados para a comparação entre universidades de dez países: Estados Unidos da América e China, que são os que apresentam as maiores produções de documentos publicados do mundo, com 173 e 165 universidades incluídas no *ranking*, respectivamente. Além desses foram incluídos Alemanha (50 universidades), França (25) e Reino Unido (45), Japão (42), Austrália (26). Para fins de comparação com o cenário do Brasil (23 universidades), incluiu-se também o conjunto de universidades da Espanha (34) e Coreia do Sul (35), pois esses dois países apresentam curvas temporais de crescimento da produção científica similares à do Brasil a partir dos anos 1980. (MANGANOTE; SCHULZ; CRUZ, 2016)

RESULTADOS: DIFERENÇAS ENTRE UNIVERSIDADES

Os indicadores de impacto recebem grande destaque nas discussões, tanto no âmbito da “cientometria leiga”, quanto sobre rankings de universidades, tornando-se balizas para metas e políticas para, por exemplo, melhorar o impacto/visibilidade por meio de colaborações internacionais. A Figura 1 mostra os indicadores de impacto (MNCS) em função da proporção de artigos em colaboração (PP(colab)) de universidades de seis países, para os quais observam-se características completamente diferentes. As universidades norte americanas apresentam uma grande dispersão para os valores de MNCS (1 indica a média mundial), com um conjunto de universidades com impacto até 50% ou mais acima da média mundial e outro grupo expressivo com impactos abaixo dessa média, mas com proporções semelhantes de colaboração. Por outro lado, a maior parte das universidades chinesas apresentam valores de MNCS próximos ou abaixo de 1, mas com uma grande dispersão no indicador de colaboração. Outros ecossistemas de universidades apresentam dispersões menores em torno de valores diferentes, sendo as francesas as mais colaborativas. É importante observar que, quanto a esse indicador de impacto, as universidades brasileiras encontram-se no mesmo intervalo que as japonesas (e sul-coreanas, não incluídas na figura).

Figura 1: MNCS versus PP(colab): universidades dos países indicados na legenda



Fonte: elaborado pelo autor baseado nos dados do *Ranking Leiden* (2019).

As colaborações internacionais, por outro lado, sugerem um quadro distinto, como ilustrado na Figura 2, que inclui todos os 10 países analisados. Considerando o conjunto completo de colaborações, as universidades do Japão e do Brasil apresentam proporções semelhantes às dos EUA e do Reino Unido. No entanto, Japão e Brasil, junto com a Coreia do Sul, apresentam percentuais de colaboração internacional bem abaixo das dos países europeus (as mais colaborativas). Os EUA apresentam proporções intermediárias e universidades chinesas registram os menores índices de internacionalização em coautorias dentre os países considerados.

Além disso, os dados apresentados na Figura 2 sugerem que colaborações internacionais, PP (colab int), correlacionam-se com os impactos das publicações mais acentuadamente do que as que se observam para PP (colab), Figura 1.

Figura 2: MNCS versus PP(colab int): universidades dos países indicados na legenda



Fonte: elaborado pelo autor baseado nos dados do *Ranking Leiden* (2019).

No entanto, os coeficientes de correlação, ρ (Pearson), variam apreciavelmente para diferentes países, Tabela 1. Para poucos casos as correlações estão no limiar de correlações fortes, para a maioria as correlações são apenas moderadas, fracas ou mesmo desprezíveis.

Tabela 1: Coeficientes de correlação de Pearson entre indicadores de impacto, MNCS, e colaboração, PP (colab) e PP (colab int)

Países	MNCS e PP(colab)	MNCS e PP(colab int)
Alemanha	0,11	0,38
Austrália	0,09	0,65
Brasil	-0,26	0,12
China	0,3	0,64
Coreia do Sul	-0,03	0,52
Espanha	0,66	0,61
Estados Unidos	0,41	0,47
França	0,69	0,69
Japão	0,29	0,52
Reino Unido	0,36	0,52

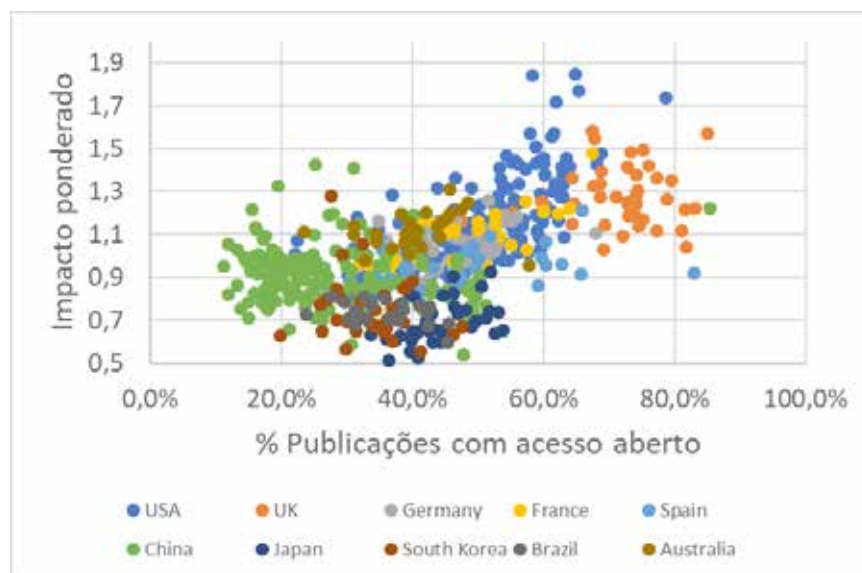
Fonte: elaborado pelo autor baseado nos dados do *Ranking Leiden* (2019).

Em geral, a Tabela 1, sugere que o impacto das publicações se correlaciona mais fortemente com as colaborações internacionais, mas esse efeito é diferente em distintos ecossistemas. Não há variação apreciável entre PP (colab) e PP (colab int) para a França, Espanha e Estados Unidos. Por outro lado, é notável para a Austrália, enquanto o efeito

das colaborações sobre o impacto das publicações é fraco para as universidades alemãs. Observa-se que para as universidades brasileiras há uma anticorrelação fraca para PP (colab) e uma correlação positiva muito fraca para PP (colab int).

Os indicadores de acesso aberto também mostram agrupamentos diferentes para universidades de países diferentes, Figura 3. As universidades chinesas apresentam os menores índices de PP (OA) (muitas apenas entre 10% e 30% do total de publicações), enquanto as do Reino Unido se concentram em torno de 75% das publicações com algum tipo de acesso aberto. As universidades de outros países estão em faixas intermediárias, sendo que as brasileiras, junto com as japonesas e sul-coreanas apresentam percentuais de publicações com acesso aberto superiores apenas às chinesas. O efeito do acesso aberto sobre as médias de citações também apresenta grandes variações entre os diferentes ecossistemas de universidades. Existe uma forte correlação positiva para as publicações das universidades francesas e estadunidenses, com coeficientes de correlação, ρ , de 0,75 e 0,7 respectivamente. A correlação é fraca ou desprezível para Espanha (0,38), Alemanha (0,36), Japão (0,35), Reino Unido (0,34), Austrália (0,2) e Coreia do Sul (0,19). Deve-se notar que no caso das universidades brasileiras e chinesas, a correlação, embora fraca, é negativa, -0,38 e -0,12, respectivamente.

Figura 3: MNCS em função de PP(AO): universidades dos países indicados na legenda



Fonte: elaborado pelo autor baseado nos dados do *Ranking Leiden* (2019).

É necessário lembrar que o efeito do acesso aberto não se limita aos pares do universo acadêmico, que são responsáveis pelas citações a trabalhos de outros autores em seus artigos. O papel é mais amplo: acesso à ciência também para um público mais amplo não

ligado às missões de ensino e extensão das universidades. No entanto, mesmo o presente recorte, limitado apenas ambiente de pesquisa propriamente dito, sugere que a resposta dos ecossistemas de universidades à prática e uso do acesso aberto é também bastante diverso.

CONSIDERAÇÕES FINAIS

Os resultados apresentados aqui indicam que a homogeneização das universidades em listas, como feito pela maioria dos rankings de universidades, esconde de fato a diversidade dos sistemas de universidades de diferentes países. Hipóteses promovidas, muitas vezes sem uma avaliação mais fundamentada, para fomentar o impacto da produção científica - colaboração internacional ou acesso aberto, os exemplos abordados nesse trabalho - não apresentam resultados com caráter universal. Um desdobramento mais aprofundado dos dados é necessário para melhor entender esses padrões. Diferenciações por contextos históricos e geográficos, como aqui exemplificados, provavelmente apresentam convoluções com as características próprias das diferentes áreas do conhecimento e os diferentes papéis das universidades em cada país. As correlações negativas entre impacto e colaboração, por exemplo, podem apontar para essas diferenças de papéis das universidades e usos da pesquisa científica. Os dados abertos do *Ranking* de Leiden permitem, portanto, informar melhor a percepção dos sistemas de universidades.

REFERÊNCIAS

HAMMARFELT, B.; RIJCKE, S.; WOUTERS, P. From Eminent Men to Excellent Universities: university rankings as calculative devices. *Minerva*, São Carlos, v. 55, n. 4, p. 391-411, 2017.

LEYDESDORFF, L.; WOUTERS, P.; BORNMANN, L. Professional and citizen bibliometrics: complementarities and ambivalences in the development and use of indicators-a state-of-the-art report. *Scientometrics*, Dordrecht, v. 109, p. 2129-2150, 2016.

LEYDESDORFF, L.; BORNMANN, L.; MINGERS, J. Statistical Significance and Effect Sizes of Differences Among Research Universities at the Level of Nations and Worldwide Based on the Leiden Rankings. *Journal of the Association for Information Science and Technology*, New York, v. 70, n. 5, p. 509-525, 2019.

MANGANOTE, E. J. T.; SCHULZ, P. A.; CRUZ, C. H. B. Effect of High Energy Physics Large Collaborations on Higher Education Institutions Citations And Rankings. *Scientometrics*, Dordrecht, v. 109, p. 813-826, 2016.

