



# Crescendo, Sobrevivendo ou Morrendo? Explorando a Comunidade dos Projetos Brasileiros no GitHub

Jailton Coelho

Instituto Federal do Paraná (IFPR) – Campus Telêmaco Borba  
Telêmaco Borba – PR, Brasil  
jailton.coelho@ifpr.edu.br

**Abstract**—O código aberto é cada vez mais importante no desenvolvimento de software, mas há dúvidas sobre sua sustentabilidade. Projetos de código aberto são geralmente gerenciados por poucos desenvolvedores, muitos deles voluntários. Neste artigo, foi avaliado a manutenibilidade de projetos de código aberto brasileiros para determinar se o ecossistema brasileiro de código aberto está crescendo, se mantendo ou diminuindo. Para isso, foi criado um conjunto de dados com os 781 projetos de código aberto brasileiros com mais estrelas no GitHub. Primeiro, foram analisadas a popularidade desses projetos em relação ao número de estrelas, observadores e bifurcações. Em seguida, foi reportado a quantidade de projetos brasileiros de código aberto no GitHub que foram criados ou ficaram sem manutenção ao longo dos anos. Por fim, foi investigado a quantidade de novos contribuidores por projeto no último ano. Os resultados mostraram que 412 (53%) projetos brasileiros estão ativamente mantidos e 101 (25%) destes projetos não tiveram novos contribuidores no último ano. Essas descobertas são importantes, pois há uma preocupação crescente com a sustentabilidade dos projetos brasileiros de código aberto.

**Keywords**—Software de Código Aberto; GitHub; Comunidade Brasileira de Código Aberto.

**Resumo**—Open source software is becoming increasingly important in software development, but there are concerns about its sustainability. Open source projects are typically managed by a small number of developers, many of whom are volunteers. In this paper, we evaluated the maintainability of Brazilian open source projects to determine if the Brazilian open source ecosystem is growing, stagnant, or declining. To do this, a dataset was created of the 781 most starred Brazilian open source projects on GitHub. First, the popularity of these projects was analyzed in relation to the number of stars, watchers, and forks. Next, we report the number of new and unmaintained Brazilian open source projects on GitHub over the years. Finally, we investigate the number of new contributors per project in the last year. Our results show that 412 (53%) Brazilian projects are actively maintained and 101 (25%) of these projects did not have new contributors in the last year. These findings are important because there is growing concern about the sustainability of Brazilian open source projects.

**Keywords**—Open Source Software; GitHub; Brazilian OSS Community.

## I. INTRODUÇÃO

Os projetos de código aberto são a base da infraestrutura digital do desenvolvimento de software moderno [1]. Em uma recente investigação conduzida pela Sonatype<sup>1</sup>, eles reportaram que a demanda por software de código aberto continua a crescer. Em 2022, por exemplo, o consumo global de software de código aberto atingiu um total estimado de 3.1 trilhões de requisições. Estes resultados, demonstram que o código aberto possui um papel fundamental para o desenvolvimento e manutenção dos softwares existentes.

O surgimento de plataformas de compartilhamento de código de software, como o GitHub<sup>2</sup>, GitLab<sup>3</sup> e Bitbucket<sup>4</sup>, contribuíram de forma significativa para impulsionar o desenvolvimento de código aberto [2]. Com 40 milhões de usuários, o GitHub é o maior repositório de código tanto para organizações privadas quanto para colaboração de código aberto. [3]. Ele hospeda importantes ferramentas como Flutter, o Microsoft's VS Code, Azure Docs, React Native do Facebook, Tensor Flow, Kubernetes, entre outros.

O Brasil também tem se destacado no cenário mundial de compartilhamento de código aberto nos últimos anos [4, 5]. Projetos brasileiros como, NICKEMANARIN/SCREENTOGIF, WIX/REACT-NATIVE-NAVIGATION e DENISIDORO/NAVI ganharam destaque no GitHub, ultrapassando a marca de 10.000 estrelas. Por esse motivo, estudos aprofundados em relação à comunidade de código aberto no Brasil precisam ser realizados para contribuir com a sustentabilidade dos atuais e o surgimento de novos projetos.

Embora centenas de projetos de código aberto brasileiros tenham sido criados nos últimos anos, muitos desses projetos também deixaram de ser mantidos pelos seus desenvolvedores. É sob esta motivação que este estudo busca responder às seguintes questões de pesquisa:

<sup>1</sup><https://www.sonatype.com/state-of-the-software-supply-chain/open-source-supply-demand-security>

<sup>2</sup><https://github.com>

<sup>3</sup><https://gitlab.com>

<sup>4</sup><https://bitbucket.org>



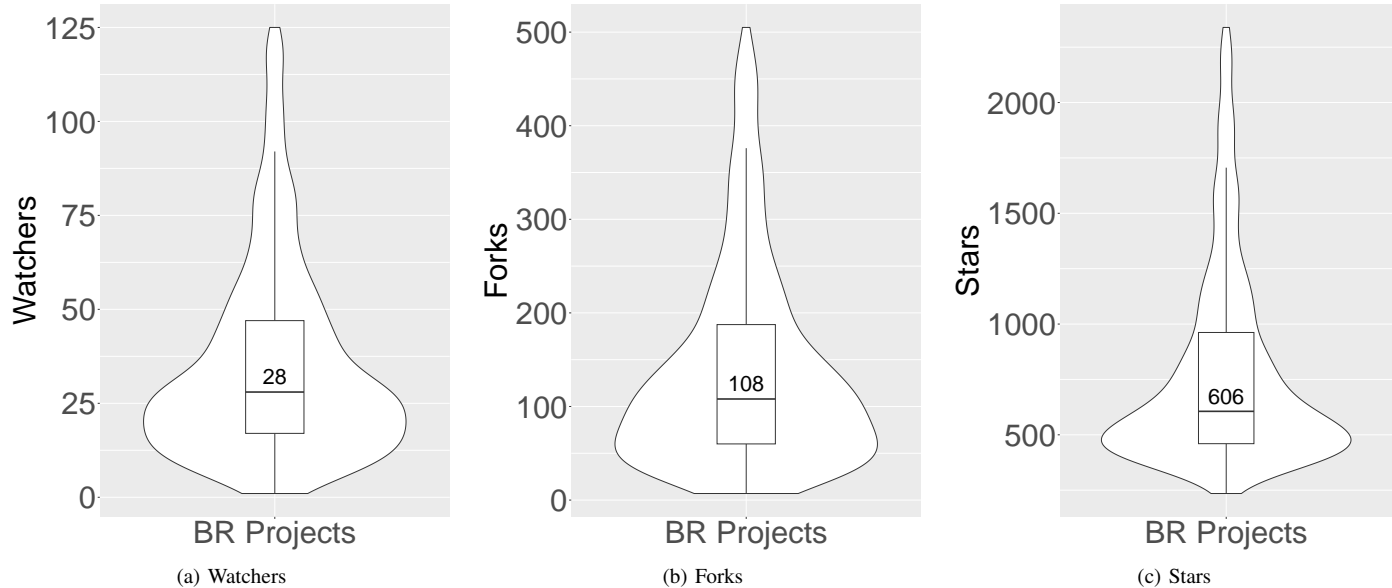


Fig. 1. Distribuição do número de observadores (*watchers*), bifurcações (*forks*) e estrelas (*stars*) dos repositórios brasileiros no GitHub, com remoção de *outliers*.

**QP1.** Qual a popularidade dos projetos de código aberto brasileiros no GitHub?

**QP2.** Quantos projetos de código aberto brasileiros no GitHub foram criados ou ficaram sem manutenção ao longo dos últimos anos?

**QP3.** Os projetos de código aberto brasileiros no GitHub continuam atraindo novos contribuidores?

Para responder estas questões de pesquisa, foram analisados 781 repositórios brasileiros hospedados no GitHub. Os resultados obtidos mostram que somente 412 (53%) repositórios tiveram atividades de manutenção no último ano e, dos 412 projetos ativamente mantidos, 101 (25%) deles não tiveram novos contribuidores no último ano. Por fim, foi constatado um aumento significativo, a partir de 2020, de repositórios que estão sem atividades de manutenção.

O restante deste artigo está organizado conforme descrito a seguir. A Seção II descreve como foi realizada a coleta das informações. A Seção III apresenta os resultados obtidos através da análise dos dados. A Seção IV apresenta as conclusões finais.

## II. PROCEDIMENTOS METODOLÓGICOS

Primeiro, foi criada uma lista dos 100.000 repositórios mais populares do GitHub, ordenados pelo número de estrelas (em fevereiro de 2023). Estrelas é um *proxy* comum para ordenar a popularidade dos projetos do GitHub [6]. O estudo foi limitado aos 100.000 repositórios mais populares para focar em projetos mais relevantes. Em seguida, foram filtrados apenas os

repositórios brasileiros desta lista, analisando a localidade do seu proprietário. Para isso, foram adotadas duas estratégias: (i) a localidade do proprietário do repositório deveria possuir alguma das seguintes *tags*: BR, Brasil ou Brazil; ou (ii) o nome da capital de algum estado brasileiro deveria estar presente na descrição da localidade do proprietário do repositório. Ao final deste processo, foi obtida uma lista de 781 repositórios brasileiros, sendo que o último repositório da lista possui 363 estrelas no GitHub.

## III. RESULTADOS

**QP1. Qual a popularidade dos projetos de código aberto brasileiros no GitHub?**

A Figura 1 apresenta a distribuição do número de observadores (*watchers*), bifurcações (*forks*) e estrelas (*stars*) dos repositórios, com remoção dos *outliers*. Os observadores (*watchers*) em um repositório GitHub é uma medida do número de usuários que expressaram interesse em acompanhar as atualizações do repositório. Bifurcações (*forks*) são cópias de um repositório GitHub que compartilha o mesmo código e configurações de visibilidade com o repositório original. Por fim, as estrelas (*stars*) são exibidas como uma medida aproximada do interesse em um repositório GitHub. Os repositórios brasileiros apresentam na mediana 28 observadores, 108 bifurcações e 606 estrelas. Os três repositórios brasileiros mais populares AVELINO/AWESOME-GO, LEONARDOMSO/33-JS-CONCEPTS e VSOUZA/AWESOME-IOS possuem 95.897, 54.826 e 41.781 estrelas, respectivamente.

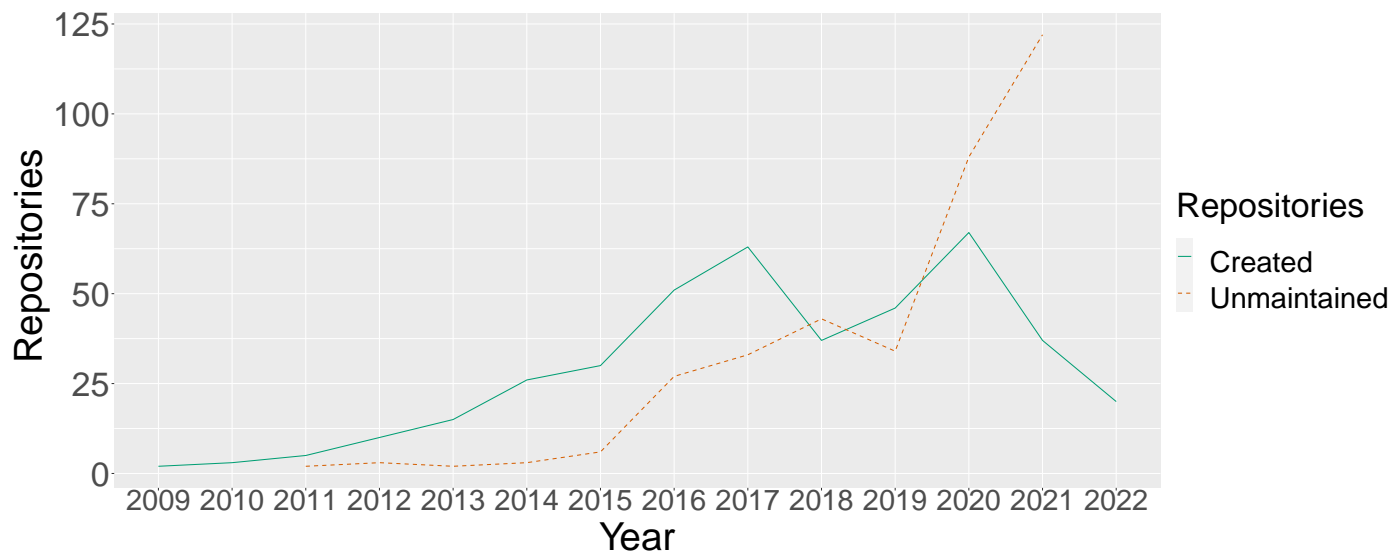


Fig. 2. Quantidade de repositórios brasileiros que foram criados (linha contínua) ou estão sem manutenção (linha tracejada) ao longo dos últimos anos.

Entretanto, estes três repositórios são classificados como listas ou livros. O quarto repositório brasileiro mais popular da lista é o NICKEMANARIN/SCREENTOGIF que possui 19.271 estrelas, tem o C# como linguagem predominante e possui um total de 96 contribuidores. Este projeto é uma ferramenta que permite gravar uma área selecionada da tela.

**QP2. Quantos projetos de código aberto brasileiros no GitHub foram criados ou ficaram sem manutenção ao longo dos últimos anos?**

Para responder esta segunda questão de pesquisa, os 781 repositórios brasileiros foram divididos em dois grupos: (i) ativos e (ii) inativos. O grupo de inativos é composto por todos os repositórios com o status arquivado (*archived*) no GitHub e também, por aqueles que apresentam indícios que deixaram de ser mantidos, como por exemplo, ficaram pelo menos um ano sem *commit*, métricas também utilizadas em outros estudos relevantes [7, 8]. O arquivamento (*archived*) é uma maneira de preservar um projeto GitHub, mas torná-lo indisponível para novas alterações. Isso pode ser útil para projetos que foram concluídos ou que não estão mais sendo ativamente desenvolvidos. Os repositórios restantes foram incluídos no grupo ativos. Ao final deste processo, o grupo ativos e inativos possuem 412 e 369 repositórios, respectivamente.

A Figura 2 mostra a quantidade de repositórios que foram criados e também, o número de repositórios que não tiveram atividades de manutenção ao longo dos últimos anos. Para o grupo de projetos ativos foi utilizada a data de criação

do repositório. Já para o grupo de projetos inativos, foi utilizada a data do último *commit*. Os repositórios que estão sem manutenção, do grupo inativos, são representados pela linha tracejada, enquanto que os repositórios ativos são representados pela linha contínua. Percebe-se que a partir de 2020, logo após o início da pandemia de COVID-19 [9] a quantidade de projetos sem manutenção superou a quantidade de projetos criados. Além disso, houve uma diminuição significativa do surgimento de novos projetos relevantes e um aumento considerável de projetos que não tiveram atividades de manutenção.

**QP3. Os projetos de código aberto brasileiros no GitHub continuam atraindo novos contribuidores?**

Para responder a essa terceira e última questão de pesquisa, foram analisados apenas os repositórios ativos, ou seja, aqueles que tiveram pelo menos um *commit* no último ano em relação a data de coleta dos dados dos repositórios. Foi considerado um novo contribuidor, apenas aqueles desenvolvedores que fizeram um *commit* no último ano, mas que nunca havia contribuído para o projeto.

A Figura 3 mostra a distribuição da quantidade de novos contribuidores dos projetos brasileiros mantidos no GitHub, cuja mediana é de apenas um contribuidor. Além disso, 101 (25%) projetos ativos não tiveram novos contribuidores. Este é um ponto que requer atenção, pois atrair e reter novos contribuidores são importantes para a sobrevivência, o sucesso a longo prazo e a continuidade dos projetos [10].

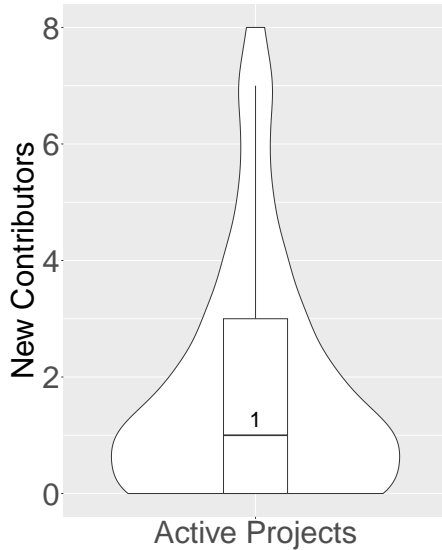


Fig. 3. Distribuição da quantidade de novos contribuidores dos projetos brasileiros mantidos no GitHub.

#### IV. CONCLUSÃO

Foram analisados 781 projetos de código aberto brasileiros com mais estrelas no GitHub. Na primeira questão de pesquisa foi apresentada a popularidade desses projetos em relação ao número de estrelas, observadores e bifurcações. Os repositórios brasileiros apresentam na mediana 28 observadores, 108 bifurcações e 606 estrelas. Estes resultados, demonstram a relevância dos repositórios brasileiros no cenário global de software de código aberto. Na segunda questão de pesquisa foi investigado a quantidade de novos projetos relevantes e também, o número de projetos que, possivelmente, deixaram de ser mantidos ao longo dos últimos anos. Foi observado que apenas 412 (53%) projetos tiveram atividades de manutenção no último ano. Além disso, constatou-se uma diminuição significativa do surgimento de novos projetos relevantes e um aumento considerável de projetos sem atividades de manutenção. Por fim, na última questão de pesquisa foi calculado a quantidade de novos contribuidores por projeto no último ano. Os resultados mostram que dos 412 projetos ativamente mantidos, 101 (25%) deles não tiveram novos contribuidores. Essas descobertas são importantes, pois há uma preocupação crescente com a sustentabilidade dos projetos brasileiros de código aberto. Como trabalho futuro, pretende-se investigar as razões que motivaram os desenvolvedores brasileiros a pararem de manter os seus projetos de código aberto.

#### REFERÊNCIAS

- [1] Nadia Eghbal. *Roads and bridges: The unseen labor behind our digital infrastructure*. Ford Foundation, 2016.
- [2] Georgios Gousios, Martin Pinzger, and Arie van Deursen. An exploratory study of the pull-based software development model. In *36th International Conference on Software Engineering (ICSE)*, pages 345–355, 2014.
- [3] Valerio Cosentino, Javier L Cánovas Izquierdo, and Jordi Cabot. A systematic mapping study of software development with github. *Ieee access*, 5:7173–7192, 2017.
- [4] Gustavo Pinto and Fernando Kamei. The census of the brazilian open-source community. In *Open Source Software: Mobile Open Source Technologies*, pages 202–211. Springer, 2014.
- [5] Marcelo Gattermann Perin, Christoph Treude, Sabrina Marczak, Leandro Melo, Igor Marques da Silva, Lucas Bibiano dos Santos, et al. A study on the geographical distribution of brazil’s prestigious software developers. *Journal of Internet Services and Applications*, 6(1):1–12, 2015.
- [6] Hudson Borges, Andre Hora, and Marco Tulio Valente. Understanding the factors that impact the popularity of GitHub repositories. In *32nd IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 334–344, 2016.
- [7] Jailton Coelho, Marco Tulio Valente, Luciana L. Silva, and Emad Shihab. Identifying unmaintained projects in GitHub. In *12th International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 1–10, 2018.
- [8] Jailton Coelho and Marco Tulio Valente. Why modern open source projects fail. In *11th Symposium on The Foundations of Software Engineering (FSE)*, pages 186–196, 2017.
- [9] Marco Ciotti, Massimo Ciccozzi, Alessandro Terrinoni, Wen-Can Jiang, Cheng-Bin Wang, and Sergio Bernardini. The covid-19 pandemic. *Critical reviews in clinical laboratory sciences*, 57(6):365–388, 2020.
- [10] Igor Steinmacher, Tayana U. Conte, Christoph Treude, and Marco A. Gerosa. Overcoming open source project entry barriers with a portal for newcomers. In *38th International Conference on Software Engineering (ICSE)*, pages 273–284, 2016.