

GitHub Business Model Analysis

ELU 605: Écologie du Logiciel

Fred Gonsalves, Jie Song

June 17, 2018

1 Introduction

The purpose of this paper is to serve as a continuation of our previous study on the intellectual property of GitHub, the largest web-based code sharing platform in the world. Since the writing of the previous paper, GitHub has officially been purchased by Microsoft for \$7.5 billion [1] and this paper will seek to understand the business model that made GitHub attractive enough of a venture for such a sale to happen.

As a quick introduction, let us discuss who GitHub is. Table 1 below gives us some of the latest statistics concerning GitHub. GitHub was founded in 2007 by three partners - Chris Wanstrath, Tom Preston-Werner, and P. J. Hyett. It was written in Ruby and was the first code sharing platform to provide distributed version control (using Git), as opposed to a centralized version control system being offered by major competitors at the time SourceForge, GoogleCode, and CodePlex. This meant that developers could clone an entire instance of a project and merge modifications much easier than ever before. GitHub's popularity grew exponentially fast as evidenced by Figure 1. In addition to using Git, GitHub is also popular today for a number of other services it provides that are discussed in the following section.

Number of users (2017)	24 million
Number of total organizations (2017)	1.5 million
Number of active repositories (since Sep, 2016)	13 million
Number of total monthly visitors (2015)	32 million

Table 1: GitHub in numbers [2].

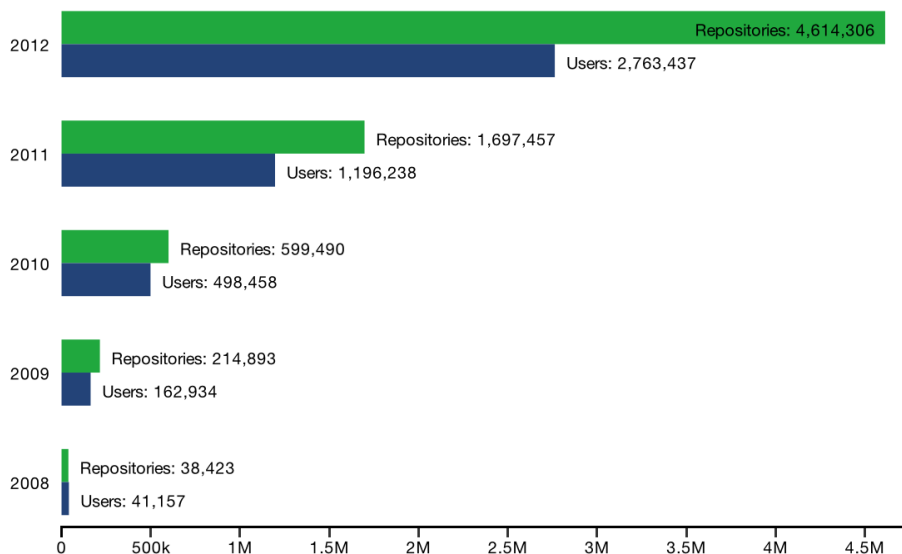


Figure 1: The growth of users and repositories of GitHub from 2008 to 2012 [3].

This paper seeks to study in detail the business model of GitHub. The paper will utilize Osterwalder's business model canvas [4] as a guide. We will start by reviewing the value proposition of GitHub, that is, what exactly GitHub has to offer in terms of services. We will then study the infrastructure that GitHub has built to provide the services it does. Next we will discuss some other options available for code sharing and how GitHub's services differ from its competitors. We will look at how GitHub has segmented the market to whom it offers its service and the various propositions it offers for each segmentation, before concluding with a short discussion on GitHub's finances.

2 Value Proposition

GitHub's core proposition is a service that allows users to host their software projects on GitHub's servers, while providing distributed version control system services of Git. GitHub and its competitors service the software development market - both open source, as well as private. This is an especially important market as the open source community grows, and as services like GitHub provide a forum for developers to showcase their talents and for software companies to headhunt. GitHub's services are provided in the form of a "Freemium" model - that is, certain services are provided for free, additional services are paid for.

2.1 Free services

Having a free account at GitHub gives users an unlimited number of free repositories hosted on GitHub servers. The only limitation is that all these repositories must be public. These repositories can be viewed by anyone, including visitors on the site. They can be forked, downloaded, and contributed to by other users. Users also have the ability to add unlimited collaborators, manage a repository's topics, and limit certain interactions with the repository [5]. Collaborators on a project may push, pull, and fork a repository. They may even manage milestones and issues, and merge or close pull requests. These public projects can have wikis, which are documents written by the owners of the repository to help new contributors understand the goal of the project, as well as get valuable technical information.

GitHub's free repositories are extremely popular among developers. GitHub currently hosts millions of free repositories, including many powerful projects like MySQL and Ruby on Rails used by GitHub themselves to power the platform.

2.2 Premium Services

In addition to all the free services mentioned above, premium users have access to a range of other services provided by GitHub. GitHub currently promotes four different options for its paid customers depending on their needs. GitHub's market segmentation strategy is discussed in detail in the Market Segmentation section. The four products, each individually catering to a specific market, include: *Developer*, *Team*, *Business*, and *Enterprise*. Table 2 below is a summary of some of the key features offered in each product.

Service	Developer	Team	Business (GitHub.com)	Enterprise
24/5 support with 8 hour response			✓	✓
Team permissions				
Organization permissions		✓	✓	✓
Fork permissions	✓	✓	✓	✓
Instance permissions				✓
Team sync				✓
Amazon Web Services				✓
Microsoft Azure				✓
Google Cloud Platform				✓
Hosted on GitHub.com	✓	✓	✓	
High availability	✓	✓	✓	user configurable
Backups	✓	✓	✓	✓
Locally controlled backups				✓
Monitoring				✓
Custom Upgrades				✓

Table 2: GitHub Product Feature Comparison [6].

GitHub’s paid options are built around a business structure that provides three core services to its customers - assurance, adaptation, and assistance - levels of which may differ from product to product. The key feature of the *Developer* pack is that it comes with unlimited private repositories. The *Team* plan offers more organizational and administrative tools for a more effective and structured collaboration. GitHub’s *Business* plan is targeted towards software development companies requiring round the clock assistance. Finally, *GitHub Enterprise* is targeted for software companies who want the ultimate control over their projects, ie. an instance of GitHub can be mounted on the company’s own private server for greater control, security, and organization of projects within the company. *Enterprise* customers may even host their repositories on Amazon Cloud Services, Google Cloud Platform, and various other cloud services supported by GitHub. Various other services are also available to *GitHub Enterprise* customers and can be found on their website [6].

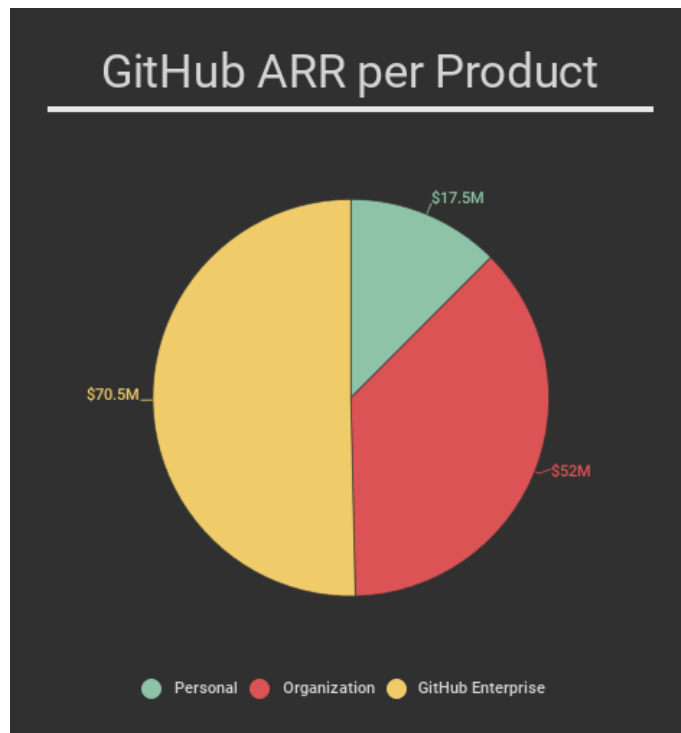


Figure 2: GitHub’s ARR per product in 2016 [7].

Apart from the features discussed above, GitHub also provides services such as documentation, issue tracking, commits history, graphs, its own text editor, PDF document viewer, and others. Figure 2 illustrates how over 50% of GitHub's revenue comes from GitHub Enterprise. According to GitHub's official website, 45% of the Fortune 100 companies now use GitHub Enterprise [2].

2.3 Marketplace

GitHub also provides a service it calls *Marketplace*. This is a place for users to share applications they create with the rest of the community either for free or as paid apps.

3 Infrastructure

In this section we will discuss the primary activities GitHub is involved in, as well as the architecture it has built in order to provide these services.

3.1 Key Activities

GitHub's key activities involve the following:

- **Repository hosting**
Although 50% of its income comes from GitHub Enterprise, which are repositories hosted on organizations' private servers, the majority of users on GitHub opt for storing their repositories on GitHub's own servers.
- **Web-based user interface**
Apart from simply hosting repositories on GitHub servers, GitHub provides a range of features on its website, which is also designed to integrate with many of the most popular integrated development environments. GitHub web pages give access to user and project statistics, graphs, easy collaboration with other users, chats, wikis, and forums.
- **Security**
To ensure code safety, compliance, and secure software development, GitHub offers security management options such as action monitoring, change verification, and permission control. This is one of the fundamental assurances given to GitHub Enterprise and Business customers. Premium customers can personalize the permissions of teams or organizations, offering better integrity of the software.
- **Support**
GitHub offers online support to all of its users, however, premium users benefit from higher levels of support. GitHub Enterprise users for example have access to 24/7 help from GitHub's support staff as well as training on GitHub's services.
- **Administration**
Administration activities include backups and upgrades.

3.2 Key Resources

- **Physical resources**
To keep the stability and performance of code deposit platform, GitHub has a strong backend architecture and server spaces. Between 2009 and 2013, GitHub used to host repositories on Rackspace [8]. In 2014, GitHub moved to the data center Carpathia [9] to expand on server space. GitHub now owns their own data centers which has helped them expand computing power and storage space in support of their exponentially growing repository base.

GitHub owns two type of facilities: Point of presence(POP) and data centers. POPs are independently connected with data centers in their respective geographic region and provide services for Internet and backbone connectivity as well as direct connect and private network interfaces to Amazon Web Services. Data centers on the other hand have three different types of cabinets - networking, computing, and storage. GitHub uses *gPanel* to manage their physical infrastructure layout, which enables developers to quickly track and locate physical

components like cabinets and PDUs, as well as easily add new components. When a project needs to expand compute or storage capacity in one of GitHub’s data centers, they will need to use an existing cabinet configuration or design a new one. As of now, GitHub stores petabytes of Git data of users of GitHub.com and has arrived at the level of around 100Gb/s across transit, Internet exchanges, and private network interfaces in order to serve thousands of requests per second.

GitHub opted to employ the use of Kubernetes, a system from Google for automating deployment, scaling and management of containerized applications. Beginning in 2017, all web and API requests to GitHub are served by containers running in Kubernetes clusters deployed on GitHub’s *metal cloud* [10]. In addition, for its *Enterprise* services, GitHub has designed custom integrations with Amazon Web Service, Google Cloud Development, and other popular cloud computing services to better serve the needs of private storage and backup.

- **Intellectual resources**

As discussed in our previous paper on GitHub’s intellectual property, GitHub makes use of a wide range of software to provide its services. Many of the core projects including Grit (Git binding software), Gollum (wiki application), and Linguist (filetype detection application) were produced in-house and are open sourced under the MIT license. GitHub benefits from open sourcing its software by sharing maintenance costs over the community, while also being able to produce robust and secure software.

In addition, the platform is also powered by some of the world’s leading open source projects. The GitHub Rails software is the website’s core architecture and was built on the open source project Ruby on Rails by Codahale. Other important open source projects that enable GitHub to provide its services are listed in Table 3.

Project	Developer
Bcrypt-ruby	codahale
ZeroClipboard	ZeroClipboard
Resque	GitHub
Swot	Lee Reilly
MySQL-Server	MySQL
Leaflet	Leaflet
Flow	Facebook
Chai	Chai.js
Primer	GitHub
ESLint	ESLint
Mocha	MochaJS
Lerna	Lerna
Linguist	GitHub
Elastic Search on Rails	Elastic
Libgit2	Libgit2

Table 3: Open source projects powering GitHub [11]

GitHub Rail and the Jobs Sinatra application are two of GitHub’s closed-core software. In addition, Github has some other closed-core software that are used to provide premium services to its business and corporation customers.

- **Human resources**

GitHub’s human resources are provided in two different ways. First, GitHub has more than eight hundred permanent employees. GitHub developers are generally tasked with the development and maintenance of certain projects which provide the cores services to GitHub and GitHub Enterprise users. Employees are also hired to provide customer support as well as training on Git and GitHub application services to GitHub Business and Enterprise users. GitHub second source of manpower comes from the open source community. GitHub itself now boasts over 28 million users, and thousands of them work on improving GitHub’s services.

In fact, many of GitHub’s current employees were formerly freelance developers on the website before getting hired by the company.

- **Financial resources**

GitHub makes an annual income of \$200 million. The financial backing of venture capitalists has helped fund GitHub’s operations immensely in terms of the number of customers it has been able to service since its conception. Section 6 discusses in greater detail the finances of GitHub, including its revenue and funding.

4 Product Differentiation

Code sharing, software development, and version control are each in their own way separate markets. If one is looking for a way to document and control revisions to code, there are multiple options available, including Git, Mercurial, CVS, SVN, Apache, and Trac, of which Git is the most popular [12]. However, using strictly a version control system like one of these does not enable a programmer to share his code. GitHub and its competitors provide a web-based platform using one or multiple of the version control software above, while also allowing projects to be hosted on their servers. This allows for projects to be collaborated and worked on by multiple developers, teams, and organizations.

GitHub’s major competitors include Bitbucket (with 6 million users as of 2017 [13], SourceForge, and more recently GitLab. But how does GitHub differ from its competitors? In this section we will look at some of the key ways in which GitHub has managed to differentiate itself from its competitors, while also comparing the features offered in the market place. Figure 3 is a table depicting the free plans offered by GitHub and some of its competitors, while Figure 4 is a comparison between GitHub services and those of its biggest direct competitor, Bitbucket.

Free Plans	Public Repos	Private Repos	Collaborators	Storage Space	Hosting	Support
GitHub Public	Unlimited	0	Unlimited	N / A	Cloud	Email / Forum
Bitbucket Small teams	Unlimited	Unlimited (1Gb /project)	5	N / A	Cloud	Email / Forum
GitLab Cloud Hosted	Unlimited	Unlimited (10Gb / Project)	Unlimited	Unlimited	Cloud	Forum
GitLab Community Edition	Unlimited	Unlimited	Unlimited	N / A	Self-hosted	Forum
Coding Free Plan	Unlimited	Unlimited	10	1GB	Cloud	Email /Forum

Figure 3: Comparison of free plans (GitHub, Bitbucket, GitLab, Coding)[14].

Feature	Bitbucket	Github
Supported VCS	Mercurial, Git	Git
Public repos	Free, unlimited	Free, unlimited
Private repos	Free up to 5 users	Starts at \$7/month for unlimited users
Integration	Jira, Crucible, Jenkins, Bamboo	Asana, Zendesk, CloudBees, Travis, CodeClimate, AWS, Windows Azure, Google Cloud, and Heroku
Popular projects hosted	Adium, Mailchimp, Opera, Python, Django	Bootstrap, Node.js, jQuery, Rails, Homebrew
Notable Extra features	Spoon, Jira integration, External authentication via Github, Twitter, Facebook, Google	Two-factor authentication, Github Pages, Github Gists

Figure 4: GitHub vs. Bitbucket [15].

One of the main selling points of GitHub that accelerated its popularity from the get-go is its employment of Git. Until this time the existing solutions, including SourceForge (the most popular code sharing platform at the time) provided solutions that did not include "distributed" version control. Linus Torvald's Git distributed version control was introduced in 2005 [16] and quickly became a very popular tool among developers, especially in the open source community. GitHub adopted Git version control since its debut in 2007, making it the first to provide code hosting using the distributed version control software. Figure 5 below shows the use of different version control systems among developers in the Eclipse Community from 2009 to 2014. it shows a clear trend line of decreasing popularity of SVN and a rising popularity of Git since 2010.

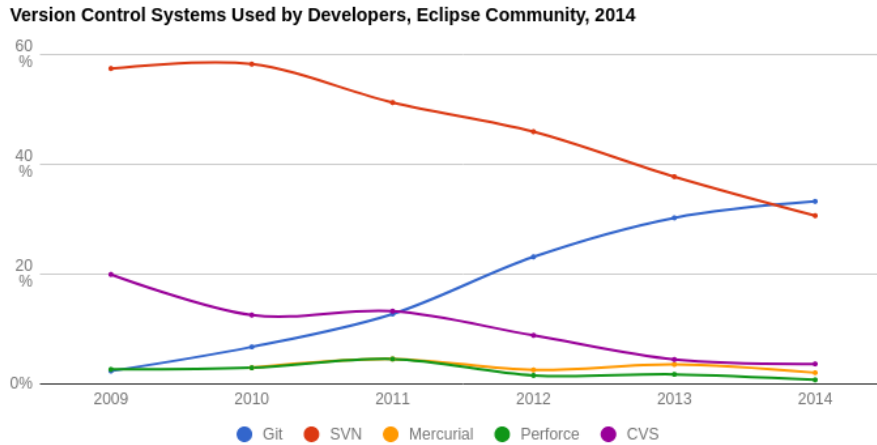


Figure 5: Version Control Systems used by Eclipse Community [17].

GitHub also does not have any advertising on its website, making it far more attractive to users, especially in its early years, when their biggest competitor was SourceForge, whose main revenue source is advertisement income. GitHub encourages a close connection between developers and project owners through its all-round social coding environment, including chats and forums. Because of the sheer number of projects and users on GitHub, developers are able to benefit from a wide variety of projects to work on, while also getting help on their own projects. As discussed in Section 3, GitHub also contributes continuously to the open source community, while also relying on the open source community. This attracts many users to the site to work on GitHub’s own projects and provides a platform for developers to showcase their talent. Every user has a public account that provide information on the user’s contributions, repositories, etc. Companies, including GitHub themselves search for and employ talented developers they find on GitHub. Because GitHub today hosts so many of the world’s biggest software projects, it has become a kind of marketplace for developers, software enterprise, and startups.

5 Market Segmentation

Although the core products of each of them may differ in many ways, GitHub and its competitors compete in a market providing software development and project management tools. While competitors Bitbucket and SourceForge segment the market based on the type of version control system used, GitHub does not do so. This is because GitHub only provides Git revision control while the others provide Mercurial, SVN, and other options. GitHub, however, segments its market based on the size of the organization owning a particular repository. This could be an individual, team (two or few individuals), small organization, or large enterprise. GitHub also segments the market into those wanting to host on GitHub’s servers to save the trouble of setting up servers, and those wanting the added security of using its own servers.

Based on these criteria for segmenting the market, GitHub provides the products discussed in Section 4. Figure 6 below is a rough positioning of GitHub’s five products on a cost-variety-personalization triangle.

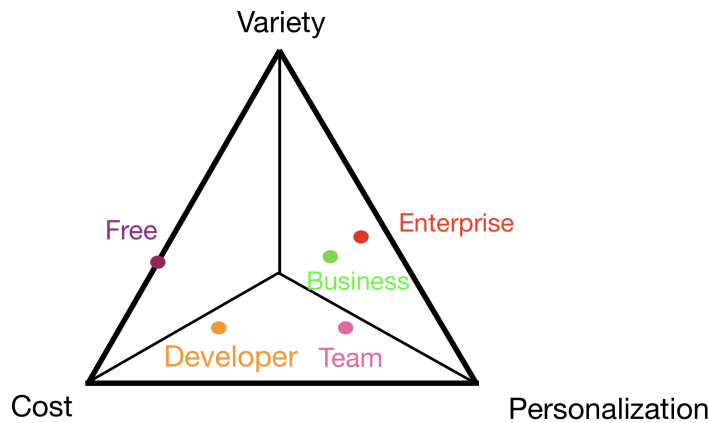


Figure 6: GitHub product characteristic distribution.

6 Finances

In this section we will gloss over the finances of GitHub. To start, let us take a quick look into the investments that have funded GitHub over the years.

GitHub was originally founded as a startup by three developers and initially served to provide enough revenue for only the three of them to live, following their exit from high-paying jobs [18]. With the rise in use of Git since 2005, GitHub found its services attracting a fast growing community of Git users. Figure 7 below depicts the history of the number of questions asked on Stack Overflow concerning the popular version control systems from 2008 to 2015. In 2012, a little over four years after its foundation, GitHub raised its first round of venture capital. At the time, valued at \$750 million, venture capital firm Andreessen Horowitz invested \$100 million, while SV Angel chipped in a smaller amount [19]. Three years later, now valued at \$1.75 billion, Sequoia Capital, Andreessen Horowitz, and two other firms invested \$250 million [20].

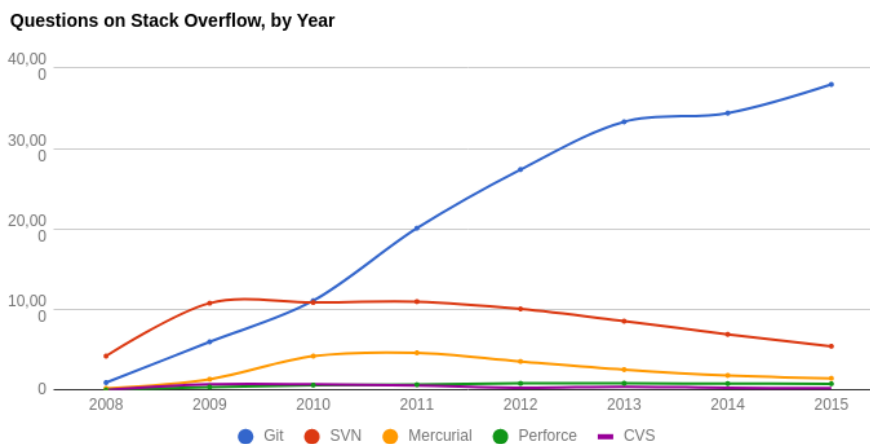


Figure 7: Stack Overflow Questions on various VCS [17].

Because of its unique service, GitHub is somewhere between a cost-driven and a value-driven product. GitHub has managed to reduce its costs significantly by open-sourcing a major portion of its product development. It's major costs right now concern primarily server space, salaries,

and rent. GitHub's exact operations cost is however, not openly available.

Owing to GitHub's commitment (at least for the moment) to be ad-free, GitHub makes the majority of its revenue from selling its product plans. GitHub's last known annual revenue, according to co-founder Preston-Werner was \$200 million [7]. As seen in Figure 2, GitHub makes over 75% of income from its business-related products. Organizations that use GitHub include AirBnB, IBM, PayPal, Spotify, and Bloomberg. Another source of GitHub's income is the Marketplace. As per the *GitHub Marketplace Developer Agreement*, GitHub takes 25% of the sales of listings in the Marketplace.

7 Conclusion

This paper has been a review on the business model that GitHub employs. GitHub is a platform that enables developers to host, control, and organize projects, promoting a stable, smart and efficient method for collaborative work. GitHub has established itself as the most used software development tool in the world over a short ten years. GitHub implements a Freemium strategy to attract as many users as possible to the platform, while providing specific, unique services to its paid clientèle to guarantee revenue flow. GitHub differentiates itself from its competitors by continuing to allow no advertising on its platform, while also providing high quality services and being a strong contributor to the open source community. All these guarantee the close connection between the company and its users, the projects and the developers, software companies and talent, and lastly ensures users' confidence in the stability and security of GitHub's service. This explains the reason of GitHub's dominance in this market.

What's more, with the acquisition of GitHub by Microsoft, in the future, GitHub could make its integration with Microsoft's public cloud service platform *Microsoft Azure*. This integration will offer developers a more complete and efficient workflow platform including code hosting, open source collaboration and environment deployment.

References

- [1] T Warren. Microsoft confirms it will acquire github for \$7.5 billion, Jun 2018. <https://www.theverge.com/2018/6/4/17422788/microsoft-github-acquisition-official-deal>.
- [2] GitHub. Github octoverse 2017. <https://octoverse.github.com/>.
- [3] B Doll. The octoverse in 2012, Dec 2012. <https://blog.github.com/2012-12-19-the-octoverse-in-2012/>.
- [4] A Osterwalder. The business model ontology: a proposition in a design science approach. Thesis.
- [5] GitHub. Github help. <https://help.github.com/articles/permission-levels-for-a-user-account-repository/>.
- [6] GitHub. Github pricing plans. <https://github.com/pricing>.
- [7] Moritz Plassnig. Github is doing much better than bloomberg thinks - here is why, Dec 2016. <https://medium.com/@moritzplassnig/github-is-doing-much-better-than-bloomberg-thinks-here-is-why-a4580b249044>.
- [8] GitHub. Github moves to rackspace. <https://blog.github.com/2009-09-15-github-is-moving-to-rackspace/>.
- [9] Data center of carpathia. <https://www.qtsdatacenters.com/>.
- [10] GitHub. Github's metal cloud. <https://githubengineering.com/githubs-metal-cloud/>.
- [11] GitHub. Projects that power github. <https://github.com/collections/projects-that-power-github>.

- [12] GitHub. Alternatives to git. <https://alternativeto.net/software/git/>.
- [13] Bitbucket. Bitbucket cloud. <https://blog.bitbucket.org/2016/09/07/bitbucket-cloud-5-million-developers-900000-teams/>.
- [14] flow.ci. Github vs. bitbucket vs. gitlab vs. coding. <https://medium.com/flow-ci/github-vs-bitbucket-vs-gitlab-vs-coding-7cf2b43888a1>.
- [15] UpGuard. Github vs. bitbucket. <https://www.upguard.com/articles/github-vs-bitbucket>.
- [16] Getting started - a short history of git. <https://git-scm.com/book/en/v2/Getting-Started-A-Short-History-of-Git>.
- [17] RhodeCode. Version control systems popularity in 2016. <https://rhodecode.com/insights/version-control-systems-2016>.
- [18] M Carney. Github ceo explains why the company took so damn long to raise venture capital. <https://pando.com/2013/06/20/github-ceo-explains-why-the-company-took-so-damn-long-to-raise-venture-capital/>.
- [19] C Schubarth. This vc firm is set to win big in microsoft's \$7.5 billion github acquisition. <https://www.bizjournals.com/sanjose/news/2018/06/04/microsoft-github-andreessen-horowitz-msft.html>.
- [20] Crunchbase. Github funding rounds. https://www.crunchbase.com/organization/github/funding_rounds/funding_rounds_list.