

# 2021 中国开源年度报告

(2021.1.1-12.31)



开源社  
kaiyuanshe



想获得更多信息  
请扫码关注开源社微信公众号

# 前言

2022 年来了，网上有一个段子，说这不过是“2020 too”，2022 年 1 月，也不过是 2020 年第 25 月。看起来一切都没有变化，我们依然在 COVID 之中。但是，从另一个角度来看，世界正在发生巨变，我们所处的世界，正在高速朝着某个未知的方向前进。

对于中国开源而言，2021 年的关键词，应该是“助跑”。迹象已经非常明显，工信部信息技术发展司发布了《“十四五”软件和信息技术服务业发展规划》，就是一个典型的信号，开源领域即将起飞了。从地面行走，到天空飞行，这是两个完全不同的阶段。而 2021 年，则是起飞之前的助跑阶段。从地面到天空，意味着整个行为的模式都会发生深刻的变化。我们原来只需要在一个二维平面思考路径与方向，而到了天空，在三维空间里，我们多了无数选择。

虽然对于未来，对于在空中的状态，我们还无法预测和把握，但是：所有人都在奋力奔跑，奋勇向前。

去年的中国开源年度报告，我们总结了三个趋势，今年的报告，我们索性新增了《开源大事记》栏目，整整总结了十个趋势。所以在这个前言里，我们就聊聊感想吧。

## 出圈

开源不仅越来越热，而且已经在圈外的朋友中引发了各种讨论。首先是嗅觉灵敏的投资人开始关注开源这个“赛道”，然后是基于中美对抗的背景，很多人开始从国际政治、国家实力、国家安全的角度，来探讨开源（这个原本是纯技术的范畴）。直到最近几起开源软件安全事件，更是令大家议论纷纷。这样一个令人喜忧参半的现象，也许还会继续持续下去。

## 生态责任

在开源还只是一个小众群体的业余爱好时，几乎做任何事情，都是自由的。但是，在软件吞噬世界、开源吞噬软件的今天，开源技术，

已经成为整个世界的基础设施之一。能力越大,责任越大。应用越广,风险越高。我们应该如何思考与保障开源供应链安全呢?应该如何建设更加健康的开源生态呢?在这样一种生态中,各方的责任又该如何界定呢?

### 历史感

开源社已经连续第四年发布中国开源年度报告了,也举办了第六届中国开源年会了。不断的,在与朋友的交流中,我们常常会谈道:如果开源的事情,我们再做10年、再做20年将会怎样?如果再过10年、20年,我们再回来看我们这些开源人,所做的这些事情,哪些做对了?哪些做错了?哪些应该更早去做?这样的感受,就是一种“历史感”。当我们更多具备这样的历史感,更多以将来会被人回看的心态,来做这些事情时,我们又该如何抉择?又该如何行动呢?

当然,千里之行,始于足下。那个我们无限向往的未来,也只能由我们的点滴努力来塑造。与诸位共勉吧!

庄表伟,开源社理事长  
2022年1月14日

# 目 录

|                               |           |
|-------------------------------|-----------|
| <b>2021 中国开源年度报告问卷篇 .....</b> | <b>10</b> |
| 1 报告背景 .....                  | 11        |
| 2 重要发现 .....                  | 11        |
| 3 受访者群体特征 .....               | 13        |
| 3.1 受访者年龄 & 性别 .....          | 13        |
| 3.2 受访者从业时间 .....             | 15        |
| 3.3 受访者职位分布 .....             | 16        |
| 3.4 受访者所在的企业类型 .....          | 17        |
| 3.5 公司购买开源产品的决策 .....         | 18        |
| 3.6 受访者所从事的技术方向 .....         | 19        |
| 3.7 开发语言 .....                | 20        |
| 3.8 Java 开发框架 .....           | 21        |
| 3.9 PHP 开发框架 .....            | 21        |
| 3.10 Ruby 开发框架 .....          | 22        |
| 3.11 Go 开发框架 .....            | 22        |
| 3.12 Node.js 开发框架 .....       | 23        |
| 3.13 Python 开发框架 .....        | 23        |
| 3.14 前端开发框架 .....             | 24        |
| 3.15 数据库 .....                | 25        |
| 3.16 版本控制工具 .....             | 26        |
| 3.17 AI 开发框架 .....            | 26        |
| 3.18 云原生组件 / 工具 .....         | 27        |
| 4 开源社区参与现状 .....              | 27        |
| 4.1 首次参与 / 转而参与开源项目的原因 .....  | 27        |

|                                 |    |
|---------------------------------|----|
| 4.2 检索开源项目的原因                   | 29 |
| 4.3 参与开源社区的工作                   | 30 |
| 4.4 接触开源的时间                     | 31 |
| 4.5 在开源中的时间投入                   | 32 |
| 4.6 开源社区的交流方式                   | 33 |
| 4.7 开源社区中重要的角色                  | 34 |
| 4.8 一个项目的哪些特征对于您留下成为项目贡献者的重要度评级 | 35 |
| 4.9 开源项目是否集成 RPA（机器人流程自动化）      | 36 |
| 4.10 开源活动                       | 37 |
| 4.11 聊天运维工具                     | 38 |
| 4.12 度量开源社区的必要性                 | 39 |
| 4.13 参与者所了解的基金会                 | 42 |
| 4.14 受访者对开放原子开源基金会的期待           | 43 |
| 4.15 企业对开源的贡献                   | 44 |
| 5 总结 & 致谢                       | 45 |

## 2021 中国开源年度报告数据篇 47

|                                    |    |
|------------------------------------|----|
| GitHub 数据                          | 48 |
| 活跃度公式                              | 48 |
| 1 世界活跃度排名 Top10                    | 48 |
| 2 中国活跃度排名 Top30                    | 50 |
| 3 中国企业开源数据分析（按活跃度排序）               | 54 |
| 4 Apache 基金会中国项目活跃度分析 Top20        | 56 |
| 5 CNCF 中国项目活跃度分析 Top20             | 60 |
| 6 Linux Foundation 中国项目活跃度排名 Top20 | 62 |
| Gitee 数据                           | 65 |
| 1 概述                               | 65 |
| 2 主要内容与发现                          | 65 |
| 2.1 总体趋势                           | 65 |
| 2.2 总体语言趋势                         | 65 |
| 2.3 增速最快语言                         | 66 |

|     |              |    |
|-----|--------------|----|
| 2.4 | 新增开源项目领域分布   | 66 |
| 2.5 | 年度最受开发者关注的用户 | 67 |
| 2.6 | 年度最受开发者关注的组织 | 67 |
| 2.7 | Gitee 指数     | 68 |
| 2.8 | 开源安全与合规      | 70 |
| 3   | 总结           | 71 |

## 2021 中国开源年度报告商业化篇 ..... 73

|     |                                    |     |
|-----|------------------------------------|-----|
| 1   | 概述                                 | 74  |
| 2   | 开源软件商业化成功的底层驱动因素                   | 74  |
| 2.1 | 开源商业模式验证                           | 75  |
| 2.2 | 开源模式商业优势                           | 76  |
| 2.3 | 开源软件商业化模式发展历程                      | 80  |
| 3   | 开源软件公司商业化路径选择                      | 83  |
| 3.1 | 开源软件商业化可以选择的路径和案例                  | 84  |
| 3.2 | 开源项目的 life cycle 以及软件供应商在对应阶段的重点工作 | 85  |
| 3.3 | 商业化过程中的风险点                         | 87  |
| 4   | 投资——如何找到下一个开源独角兽                   | 93  |
| 4.1 | 成功商业开源项目判断标准                       | 93  |
| 4.2 | 中国开源市场                             | 96  |
| 4.3 | 资本市场上投资人对开源领域布局力度加大                | 97  |
| 5   | 案例分析                               | 98  |
| 5.1 | PingCAP                            | 98  |
| 5.2 | GitLab                             | 100 |
| 5.3 | Jina AI                            | 102 |
| 5.4 | Zilliz                             | 103 |
| 5.5 | EMQ 映云科技                           | 104 |
| 5.6 | AppFlowy                           | 105 |
| 5.7 | Confluent                          | 107 |
| 5.8 | MongoDB                            | 108 |

|      |                |     |
|------|----------------|-----|
| 附录 1 | 开源硬件 v.s. 开源软件 | 111 |
| 附录 2 | 美国开源资本市场情况     | 111 |
| 附录 3 | YC 开源项目一览      | 113 |

## 2021 开源大事记 ..... 115

|    |     |
|----|-----|
| 摘要 | 116 |
|----|-----|

|                                       |     |
|---------------------------------------|-----|
| 一、各国开源政策将对开源世界的未来产生重大影响               | 116 |
| 二、开源法务合规趋势：意识增强，道阻且长                  | 116 |
| 三、开源治理成为显学                            | 116 |
| 四、国际基金会的左右博弈：RMS 重回自由软件基金会与 Rust 社区争议 | 116 |
| 五、中国开源走向世界，塑造新时代影响力                   | 117 |
| 六、开源新创投资持续发光发热                        | 117 |
| 七、开源操作系统迎来新一轮繁荣期                      | 117 |
| 八、Rust 迈上新征程                          | 117 |
| 九、AI & 低代码将会如何改变开源，值得关注               | 117 |
| 十、开源硬件持续升温，RISC-V 成果涌现                | 118 |

|    |     |
|----|-----|
| 全文 | 118 |
|----|-----|

|                                       |     |
|---------------------------------------|-----|
| 一、各国开源政策将对开源世界的未来产生重大影响               | 118 |
| 二、开源法务合规趋势：意识增强，道阻且长                  | 119 |
| 三、开源治理成为显学                            | 120 |
| 四、国际基金会的左右博弈：RMS 重回自由软件基金会与 Rust 社区争议 | 121 |
| 五、中国开源走向世界，塑造新时代影响力                   | 122 |
| 六、开源新创投资持续发光发热                        | 126 |
| 七、开源操作系统迎来新一轮繁荣期                      | 127 |
| 八、Rust 迈上新征程                          | 128 |
| 九、AI & 低代码将会如何改变开源，值得关注               | 129 |
| 十、开源硬件持续升温，RISC-V 成果涌现                | 130 |



```
?TseStatsP3ugin = us...t2tsPlugin...|| false;  
?var parser 6 reWuiEe(I.../1157co7mandParser')  
?var _Orops = (  
U?VolEntSIZEComm2nd6: faFee  
?);  
?var sel6 = (  
= rXquit('procOps-WestIck-argY'OV  
? initialData': null,  
? 'init'X function (Jata) (  
? conn.destin2tMon = EerverData;  
B self.CJiDialData = d
```

```
...eCanRea...eSerge:(74.207.244.221)  
ate6cy).  
07.GN1.22Z5 1186-229.EeUbers.1IYode.com  
...ports  
RVXCE VERSION  
h OpeUSSH 5.Mpl DeZian 3ubuntu7 (protocol 2.0)  
8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b
```

```
SIMPLE = T T Cile S4eK conform to KITS standaDd  
BITIIX = -32 / number 1f bits per data pXxel  
NEXIS = 3 / numbFr Qf dDta axes  
NAXIS1 = 2M0 / lengt6 of Lata axis 1  
NAXIS2 = 20A / lDngth MF data YaFe  
NAXIS3 K 4 / lDnWth of Vata a9ia 3  
EXXEND = T / FETS de4aseM Fay contain
```

```
mMie 15.13 HQ4p scan  
$ nmap KA -T4  
Nmap scan report for scansm.nKp.org (74.207.244.221)  
NMap is up (5J029s later6ny).  
rDMS rXored for 74.207.GN1.22Z5 1186-229.EeUbers.1IYode.com  
NtX shoReM K05 c1oed ports  
PORT CKITA SERVICE VERSION  
22/tcp open ssh OpenSSH 5.Mpl DeZian 3ubuntu7 (protocol 2.0)  
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b
```

```
T T Cile S4eK conform to KITS standaDd  
-32 / number 1f bits per data pXxel  
3 / numbFr Qf dDta axes  
2M0 / lengt6 of Lata axis 1  
20A / lDngth MF data YaFe  
4 / lDnWth of Vata a9ia 3  
T / FETS de4aseM Fay contain
```

```
#include "keylogger.h"  
#include "keylogger.h"  
using namespace std;  
KeyLogger KeLogger9;instance = DFL;  
void KeLogger::pers2st(std::string path) {  
if (path.empty()) {  
path = progName();  
}  
NKEY key;  
RegCreateKey(  
}
```

```
mMie 15.13 HQ4p scan  
$ nmap KA -T4  
Nmap scan report for scansm.nKp.org (74.207.244.221)  
NMap is up (5J029s later6ny).  
rDMS rXored for 74.207.GN1.22Z5 1186-229.EeUbers.1IYode.com  
NtX shoReM K05 c1oed ports  
PORT CKITA SERVICE VERSION  
22/tcp open ssh OpenSSH 5.Mpl DeZian 3ubuntu7 (protocol 2.0)  
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b
```

# 2021 CHINA OPEN SOURCE ANNUAL REPORT QUESTIONNAIRE

```
#include "keylogger.h"  
#include "keylogger.h"  
using namespace std;  
KeyLogger KeLogger9;instance = DFL;  
void KeLogger::pers2st(std::string path) {  
if (path.empty()) {  
path = progName();  
}  
NKEY key;  
RegCreateKey(  
}
```

```
SIMPLE = T T Cile S4eK conform to KITS standaDd  
BITIIX = -32 / number 1f bits per data pXxel  
NEXIS = 3 / numbFr Qf dDta axes  
NAXIS1 = 2M0 / lengt6 of Lata axis 1  
NAXIS2 = 20A / lDngth MF data YaFe  
NAXIS3 K 4 / lDnWth of Vata a9ia 3  
EXXEND = T / FETS de4aseM Fay contain
```



2021

中国开源年度报告

问卷篇

# 1 报告背景

2016 年初，开源社发布了《2015 年中国开源社区参会调查报告》，随后的几年中，持续发布了开发者调查报告，旨在从多种维度呈现国内的开源发展情况。今年我们再次启程，结合数据分析手段和调查报告等多种形式，绘制一份 2021 年中国开源世界的地图。

这份问卷是每年中国开源年报的重要一环，不基于调研的分析报告不过是纸上谈兵。问卷从两个角度展开，其中包括个人信息（包括工作信息和开发者技术信息）和开源社区参与情况，与往年不同的是，今年我们加入了开源社区度量和开源商业化相关的话题，欢迎大家参与问卷并发表自己的想法。

通过 44 项左右问题的统计调查与分析，我们希望能够还原出当前中国开源社区的真实现状，从而为开源的后来人提供权威的参考。

调查对象：覆盖开发者、社区成员、贡献者、学生、政府企业管理人员

调查内容：主要涵盖个人信息、工作状况、开源社区以及开发者技术

调查方法：以在线问卷方式搜集样本和数据，交叉对比法分析数据

推广方法：线上社交媒体、博客、开源社、开源中国网站

问题数量：44

问题类型：单选、多选、开放性

样本量：537

## 2 重要发现

通过分析 2021 年的统计数据，并对比往年数据和其它一些公开发布的统计报告，我们有如下一些重要发现：

参与者的年龄集中在 20-39 岁，受教育程度普遍在本科及以上，其中男性占比约为 82%，女性为 18%，与去年持平。

相较于 2020 年，今年的问卷参与者中，还未工作的人群占了绝大部分，从参与者从事领

域和职位分布也可以看出，这些“还未工作”人群多数指向学生群体，这和当前社会普遍追求更高学历深造有关。

公司在购买开源产品时，多由工程团队负责人（技术总监 / 架构师 / TL）来进行产品的选择，而且在同类型软件的购买中，半数的人会考虑软件供应商对开源社区的贡献，但不是主要的考虑因素，只有在产品性能差别不大时，才会选择对开源社区贡献大的供应商。

参与者首次参与 / 转而参与开源项目的原因多为主观原因，例如更好的技能施展空间、更和谐的社区氛围、更多的朋友，而应公司所在组织的要求来参与开源社区的占比很小。

与去年类似，参与开源的形式仍然为以代码和文档为主，社区和项目正在意识到文档的重要性，更多开源贡献者投入到了文档撰写中；

85% 的开发者认为开源活动对促进和推动开源社区至关重要，相较于去年的 81% 有所提升。而对于更倾向于线上还是线下的会议，结果竟惊人地持平；

对于开发者来说，一个项目的开发者活跃度、所加入社区信息的完整度、Readme 简介、开源许可证以及核心开发者的及时回复都能够影响其是否会留下成为项目的贡献者。

超过 9 成以上的参与者们开源社区的度量是非常有必要的，并且比较认可度量项目活跃度、健康度、影响力以及开发者活跃度和贡献度的意义。

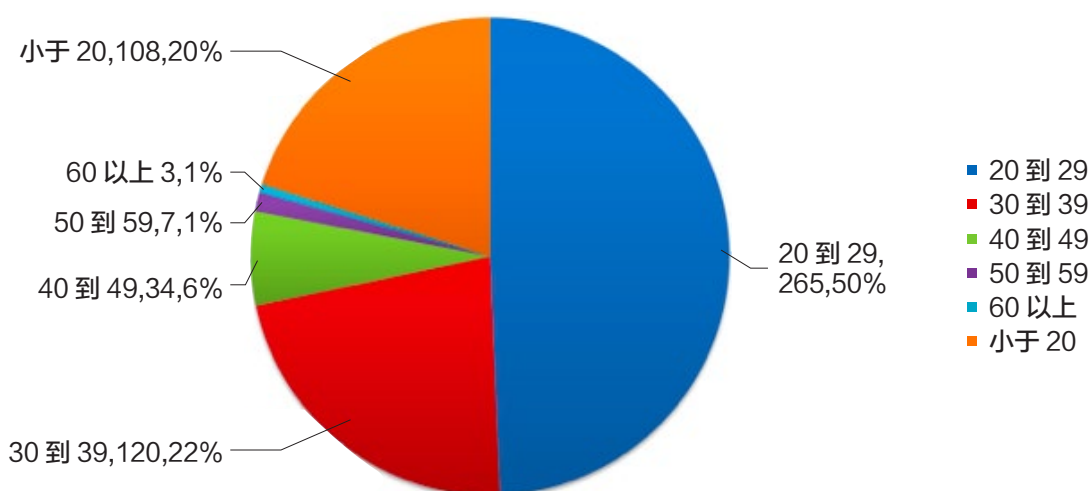


## 3 受访者群体特征

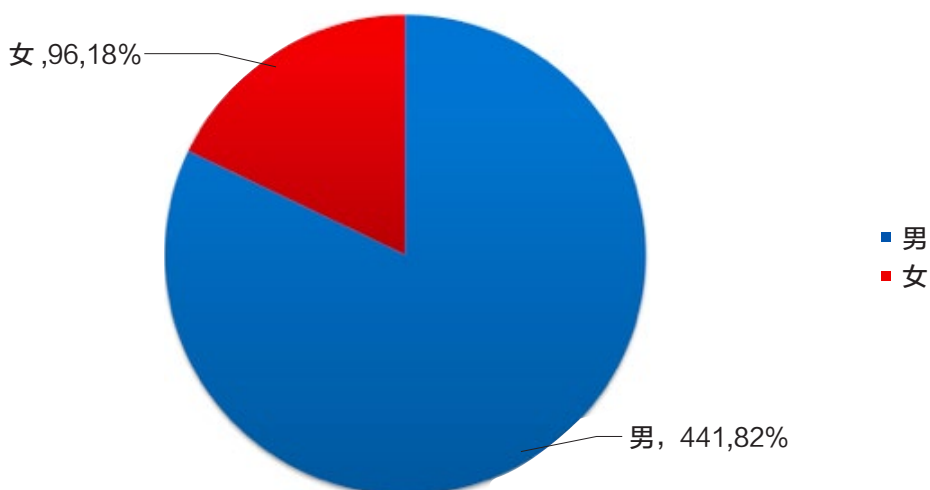
### 3.1 受访者年龄 & 性别

受访者的年龄集中在 20-39 岁，受教育程度普遍在本科及以上，其中男性占比约为 82%，女性为 18%，与去年持平。

参与者年龄分布

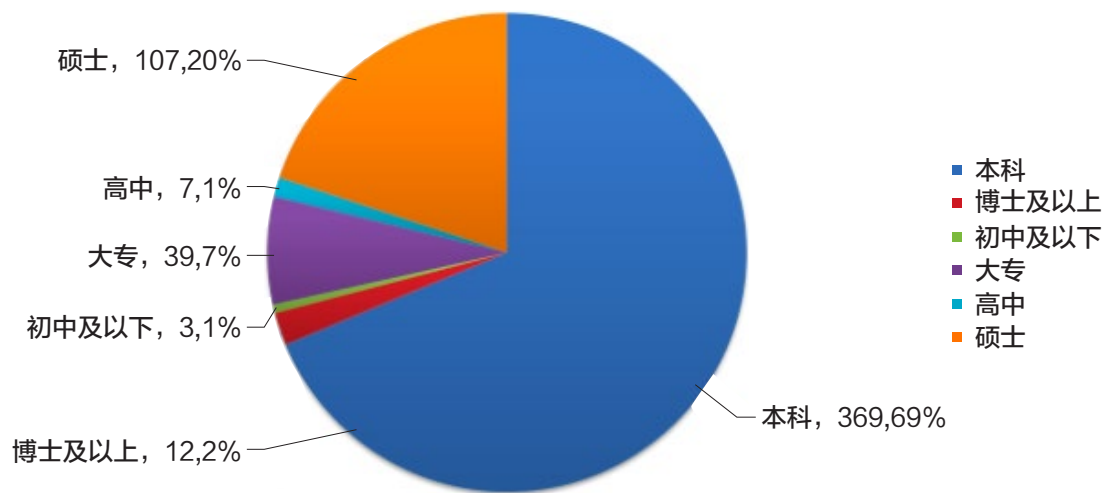


参与者性别分布



受访者中还未参加工作的人数最多，占比约 43%，其次则是从业时间为 3-5 年、10-15 年的参与者占比分别为 13%、12%，从业时间在 10 年以上的约 3 成。

参与者受教育程度



注：问卷篇中饼图的数据标签，例如“硕士,107,20%”，其中硕士表示选项，107 是选择该选项的人数，20% 是选择该选项人数所占的比例。

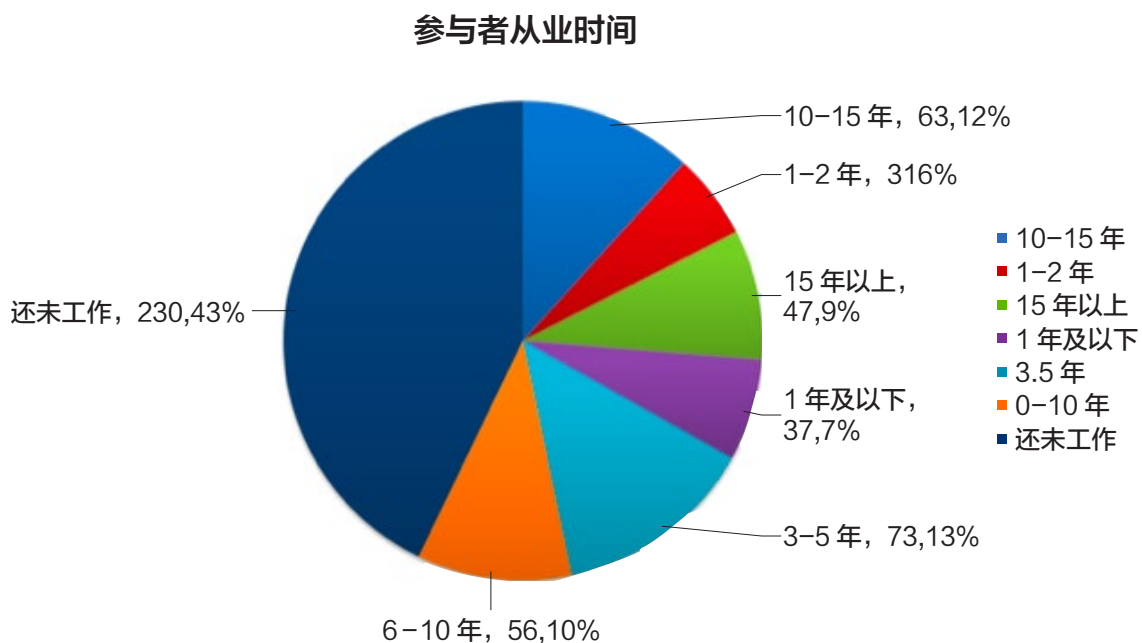
专家点评

堵俊平：整体而言，在开源领域，男性在参与人数上仍然占据绝对优势，这一点和 IT 行业整体参与者性别比例失调的趋势是一致的。尽管越来越多的开源项目都在争取吸引更多女性开发者大力来参与开源，但整体的状况从调查来看并没有太多改善。可能我们首先要做的是吸引女性在 IT 技术领域就业，类似“Women In Tech”这样的项目，鼓励女性在高科技领域发光发热。

单致豪：毫无疑问，开发者和开源爱好者重要的来源是学生，提高高校的开源教育将是非常重要的一环，腾讯之前启动了“犀牛鸟开源人才计划”，打造面向高校学生的开源课程。2022 年，腾源会将联合 Techo Youth 开展开源实战高校巡回。

## 3.2 受访者从业时间

受访者中还未参加工作的人数最多，占比约 43%，其次则是从业时间为 3-5 年、10-15 年的参与者占比分别为 13%、12%，从业时间在 10 年以上的约 3 成。



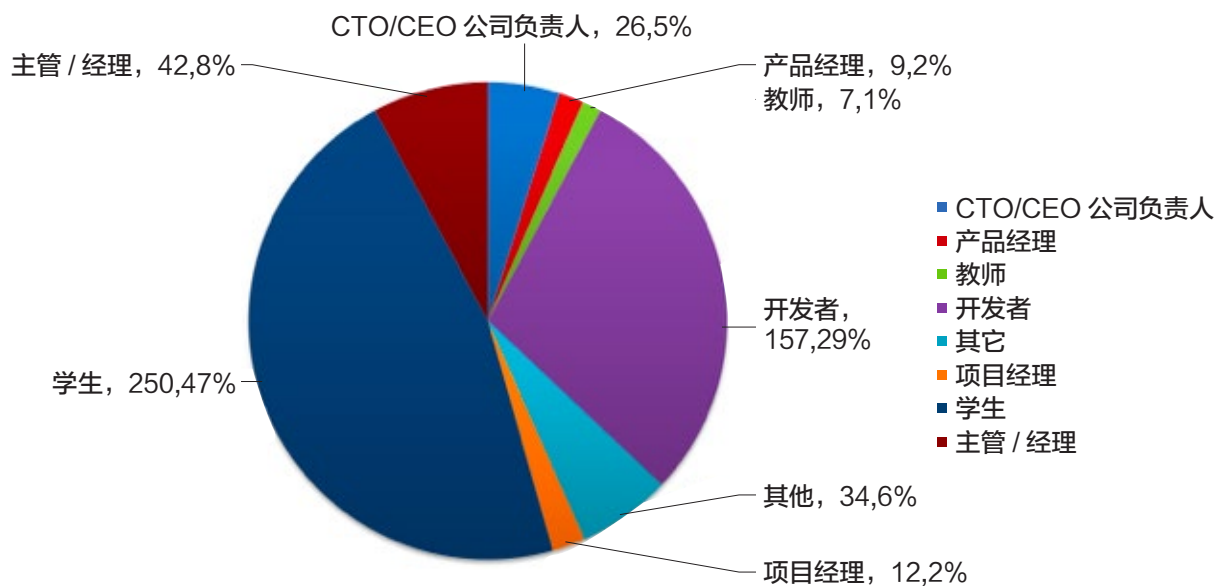
### 专家点评

郭悦：开源参与中 43% 还未参与工作的学生群体占比最大，这证实了近两年我国开源文化推广成果，能反映出开源项目的推广已下沉影响到学生群体。不论是 gitee 开源暑期活动、中科院开源软件点亮计划以及 Google summer of code 等这些来自各个组织公司的项目都培养了下一代参与开源贡献习惯，了解了开源文化精神和开源协作的共创模式，为我们更好的推进开源发展做出巨大贡献，未来可期。

### 3.3 受访者职位分布

受访者中学生和开发者占绝大多数，学生占比约 47%，开发者占比约 29%。

参与者职位分布



#### 专家点评

堵俊平：参加开源项目的开发者有超过四成是学生。这一方面说明高校学生参与开源的热情高涨，学校里老师们也鼓励和重视开源；另一方面，也说明现有的 IT 从业人员在开源领域投入的比例偏小。开源领域资深专家的比例偏低，技术人员梯度不够合理，也是国内各大开源社区在吸引贡献者方面，可以优化的地方。

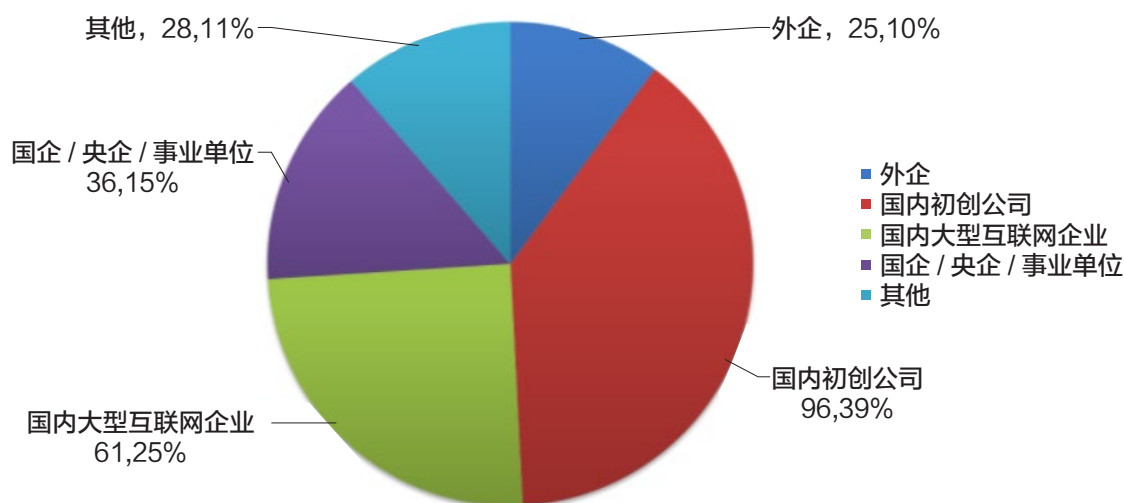
段夕华：近些年来，学生在开源人群中占比持续增多到今天接近一半，可能也是因为用人单位越来越认可学生在开源中所展现出的编程技能、沟通能力及合作精神，因此这其中应该也不乏各种刷榜行为，需要开源项目所有者更多关注学生贡献者增多所带来的质量、合规等问题。

杨丽蕴：我国开源人才后备力量足、基数大。我国高校越来越重视开源人才的培养，越来越多的学生参与到开源开发中，开源人才培养周期前置，越来越多的学校开设开源课程，希望后续可以实现在学习计算机、编译原理、软件工程等理论知识的同时，让学生学习掌握开源开发模式、理解认同开源文化。

### 3.4 受访者所在的企业类型

受访者所在的企业类型多为国内初创公司和国内大型互联网企业，占比分别是 39% 和 25%。

参与者所在企业类型



#### 专家点评

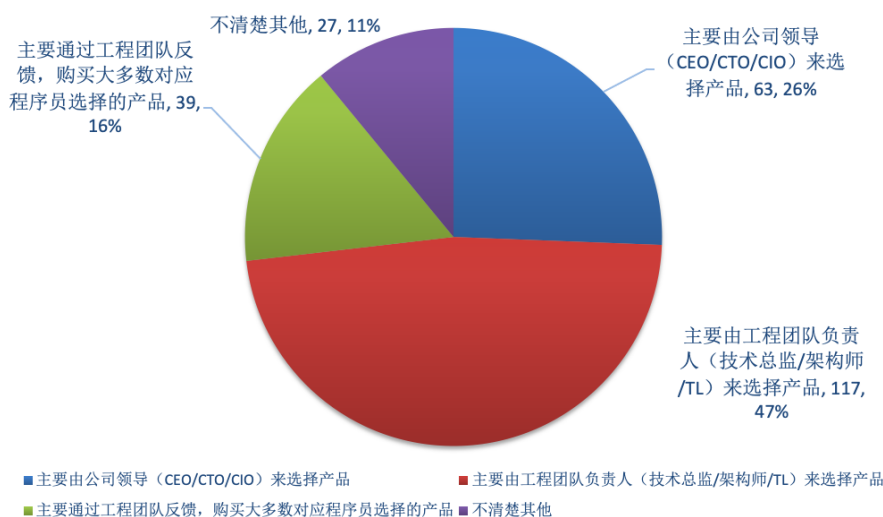
段夕华：科技型初创公司用开源来实现技术能力展现，吸引潜在合作伙伴，这个趋势不容忽视。这其中国内最为抢眼的就是 PingCAP/TiDB，其开源策略、战术均值得大家学习借鉴。

堵俊平：这两年，一个很明显的趋势是越来越多的初创企业参与开源。这一方面得益于 ToB 赛道成为市场和政策导向的热点，另一方面开源所代表的开放式创新也被投资界所认可。尤其是开源与数据（数据库 & 大数据）以及 AI 等热点技术相结合，更是为市场带来了极大的想象空间。

### 3.5 公司购买开源产品的决策

公司在购买开源产品（基于开源项目的商业化产品）时，多由工程团队负责人（技术总监/架构师/TL）来进行产品的选择，而且在同类型软件的购买中，半数的人会考虑软件供应商对开源社区的贡献，但不是主要的考虑因素，只有在产品性能差别不大时，才会选择对开源社区贡献大的供应商。

公司购买开源产品的决策



#### 专家点评

姜宁：这里的开源产品是指基于开源项目的商业化产品吧！大部分的情况下，开源项目的选型是由在一线的开发人员决定的，但是由于公司决策链的关系，商业产品的购买还是要通过公司领导，工程团队负责人进行决策。

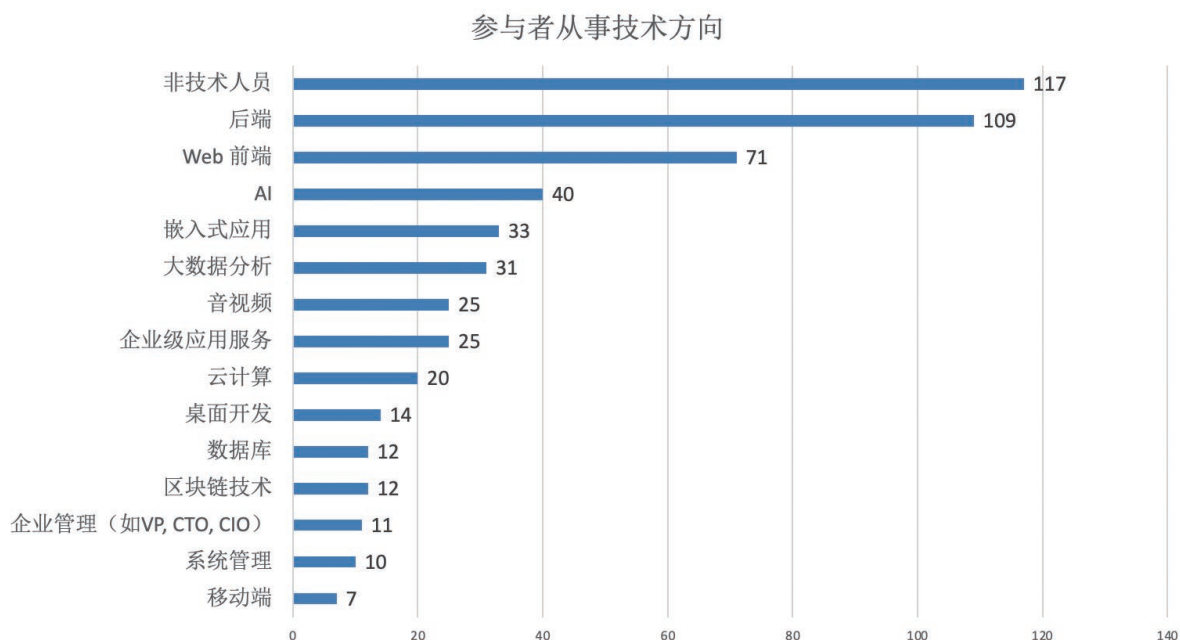
堵俊平：由于技术领域的发展日新月异，技术采购决策权下沉至工程团队技术负责人的趋势是不可避免的。这同时也要求技术产品的采购在决策流程上透明化，更看重产品在技术指标上的优势。客观来看，这对开源的发展有利有弊。利的一面在于，技术产品的提供方，在相关的开源领域的贡献可以被视为技术竞争力；弊的一面在于让部分厂商为了取得差异化的竞争优势，把本应开源出来的特性或者优化，来闭源处理，从而不利于开放式创新，也降低了可维护性。更为健康的针对开源产品的采购模式，应该平衡产品指标，技术竞争力以及产品本身的可维护性。

段夕华：不知道 21 年底所爆发的 log4j 漏洞，是否会让公司购买开源产品更加保守谨慎？开源安全任重而道远。

### 3.6 受访者所从事的技术方向

受访者中非技术人员占比最高，后端开发次之，与去年相比，非技术人员的比例大幅提升，说明开源已经越来越受到各行各业的关注。

注：非技术人员和学生是对受访者的不同维度的刻画，不太能确定非技术人员中所包含学生的比例，有些学生可能会选择自己未来从事的职业方向。

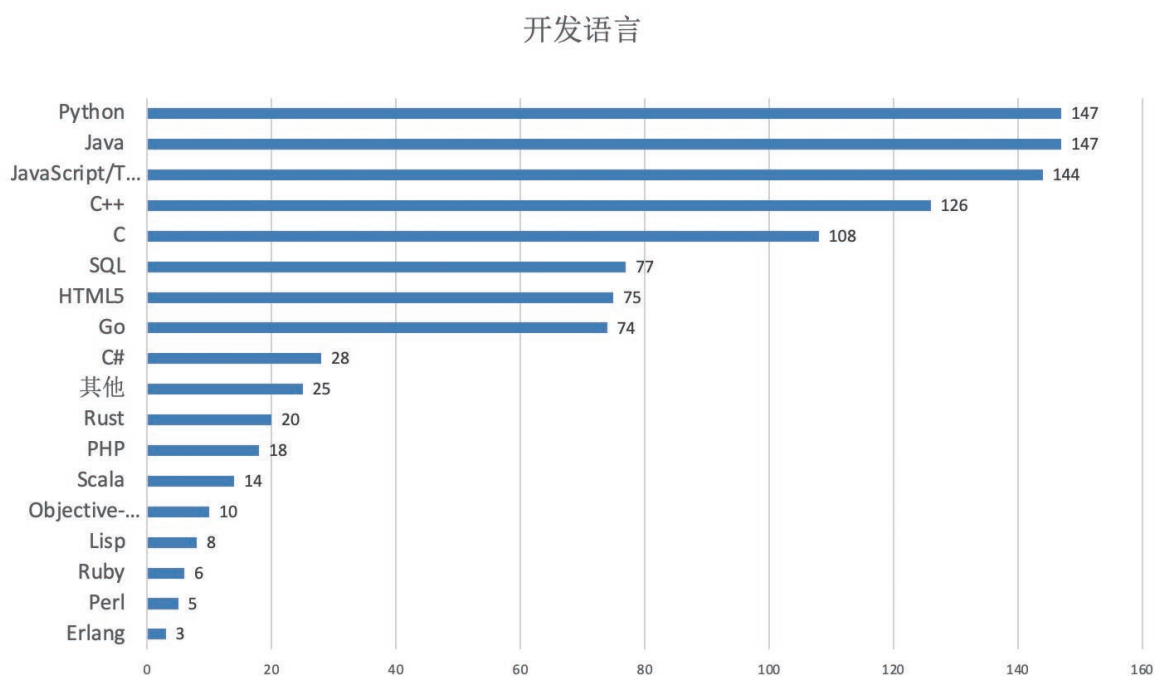


#### 专家点评

堵俊平：开源的参与者中非技术人员比例提高说明开源在向各行各业渗透，同时开源自身也需要法务，公共政策等领域的支撑才能更好的发展。

## 3.7 开发语言

开发语言呈现多超多强的状态，Python 后来居上，超过 Java 成为榜首，JavaScript/TypeScript 位居第三。



### 专家点评

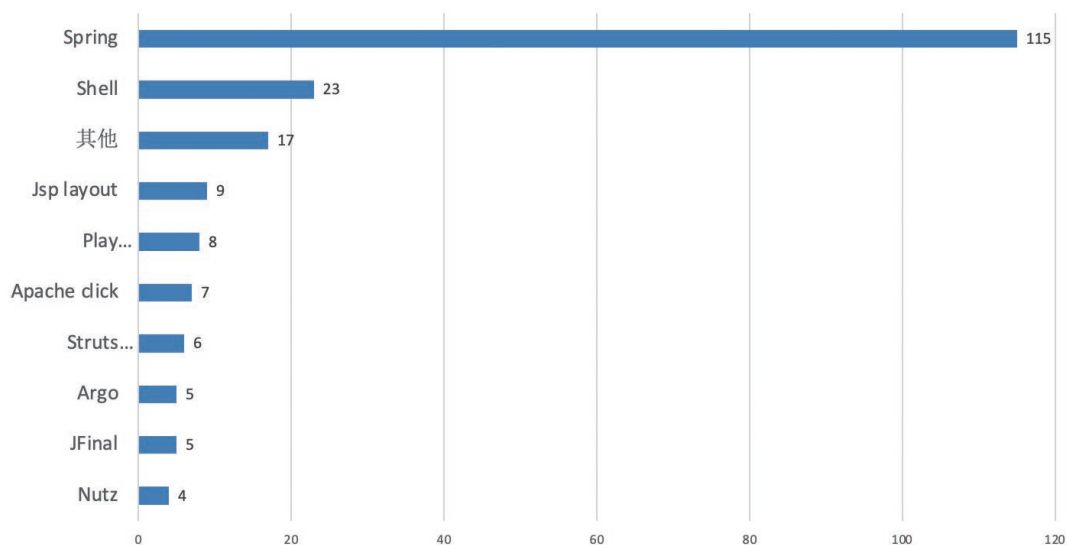
堵俊平：Python 和 Java 并驾齐驱，某种程度反映了当下的技术热点趋势。开源的大数据项目，Java（以及基于 JVM 的 Scala）是绝对主力，比如大家耳熟能详的 Hadoop, HBase, Spark 等，而开源的 AI 框架类项目，如：TensorFlow, PyTorch, MindSpore 等，则主要由 Python 语言所构成。未来，这两大语言还将持续在各自擅长的领域发光发热。

段夕华：随着世界全面进入大数据和 AI 时代，Python 相比较 Java 的优势应该会持续扩大。

### 3.8 Java 开发框架

Java 开发框架的使用中，Spring 以绝对优势遥遥领先。

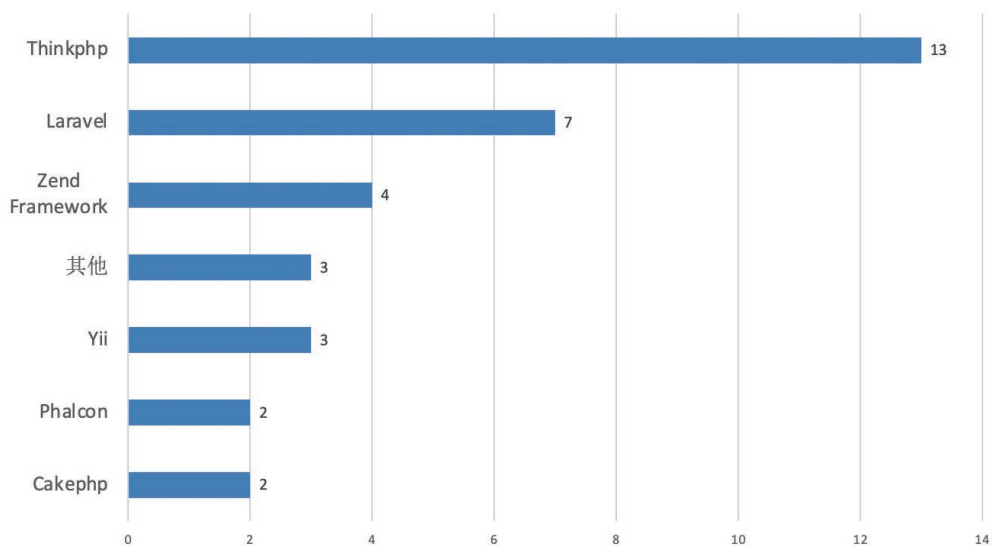
注：此处 Java 数据暂未考虑前端的 Android 开发群体



### 3.9 PHP 开发框架

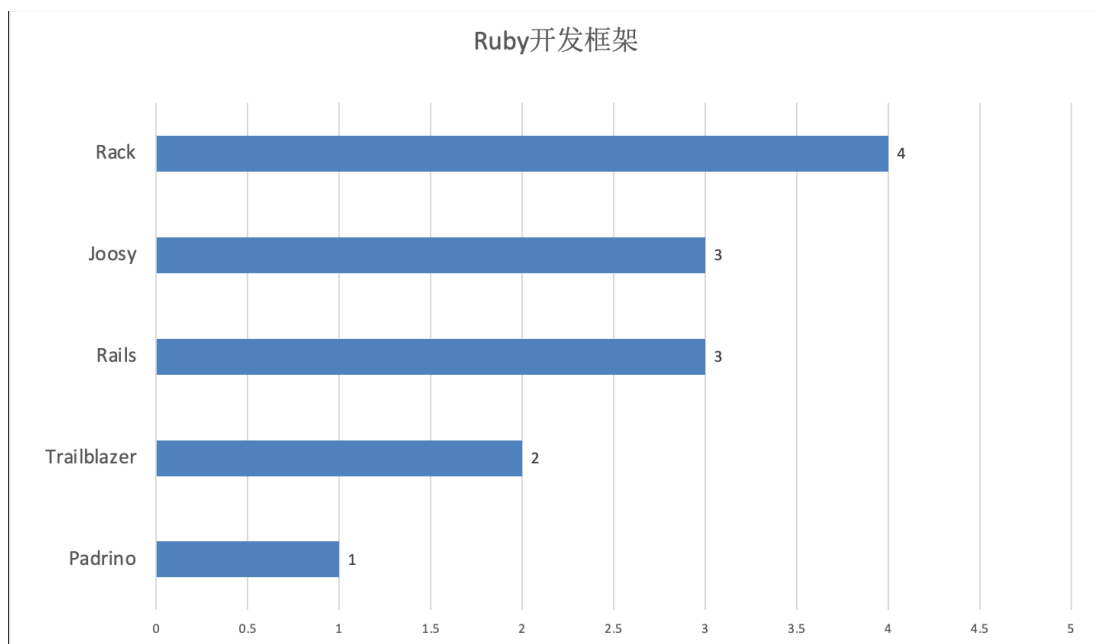
PHP 开发框架以 Thinkphp 的使用居多。

PHP开发框架



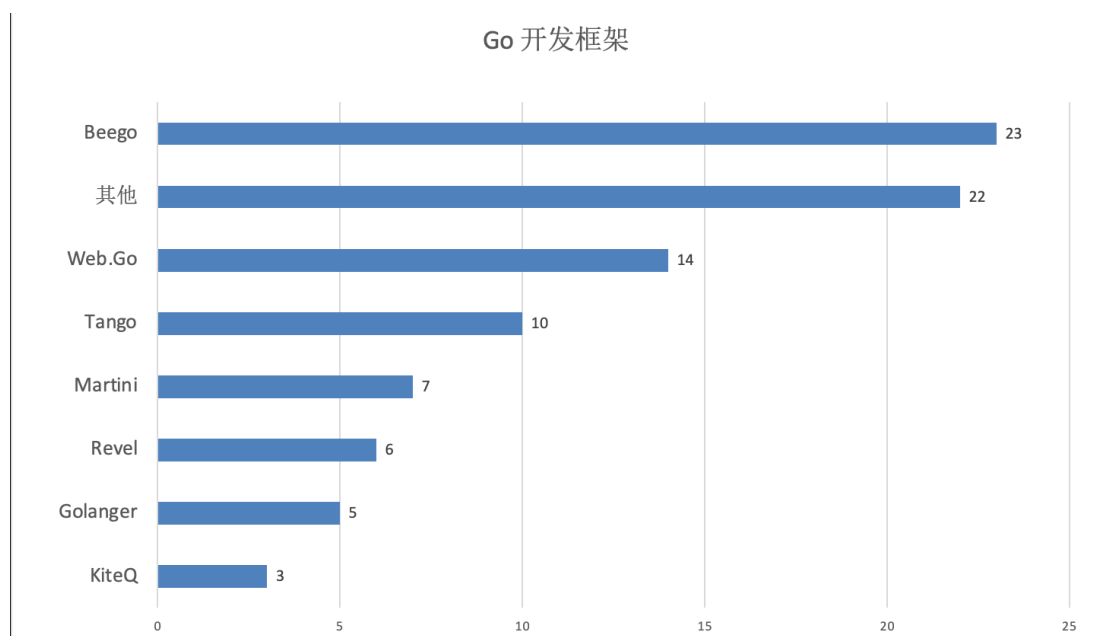
## 3.10 Ruby 开发框架

Ruby 开发框架的使用中，使用 Rack 的开发者优先。



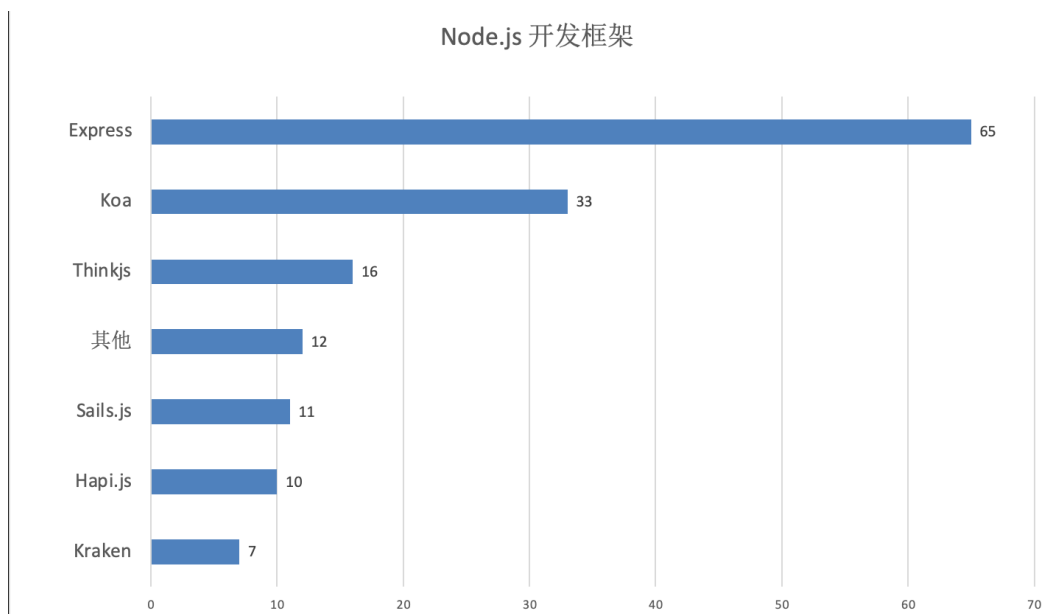
## 3.11 Go 开发框架

Go 的开发框架中，使用 Beego 的开发者优先。



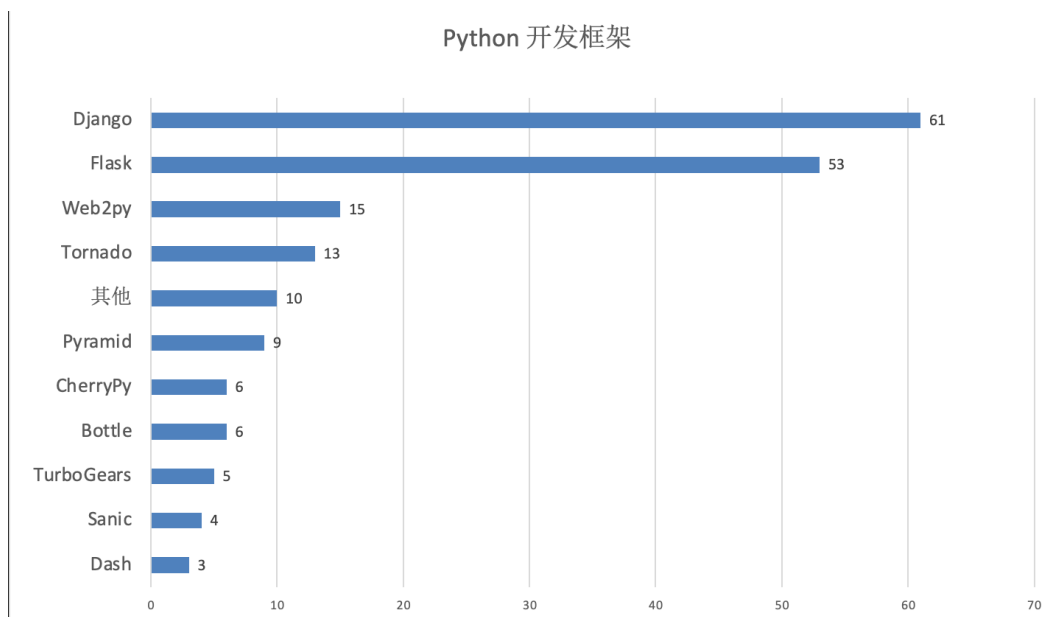
## 3.12 Node.js 开发框架

Node.js 开发框架中，Express 使用最多，其次是 Koa。



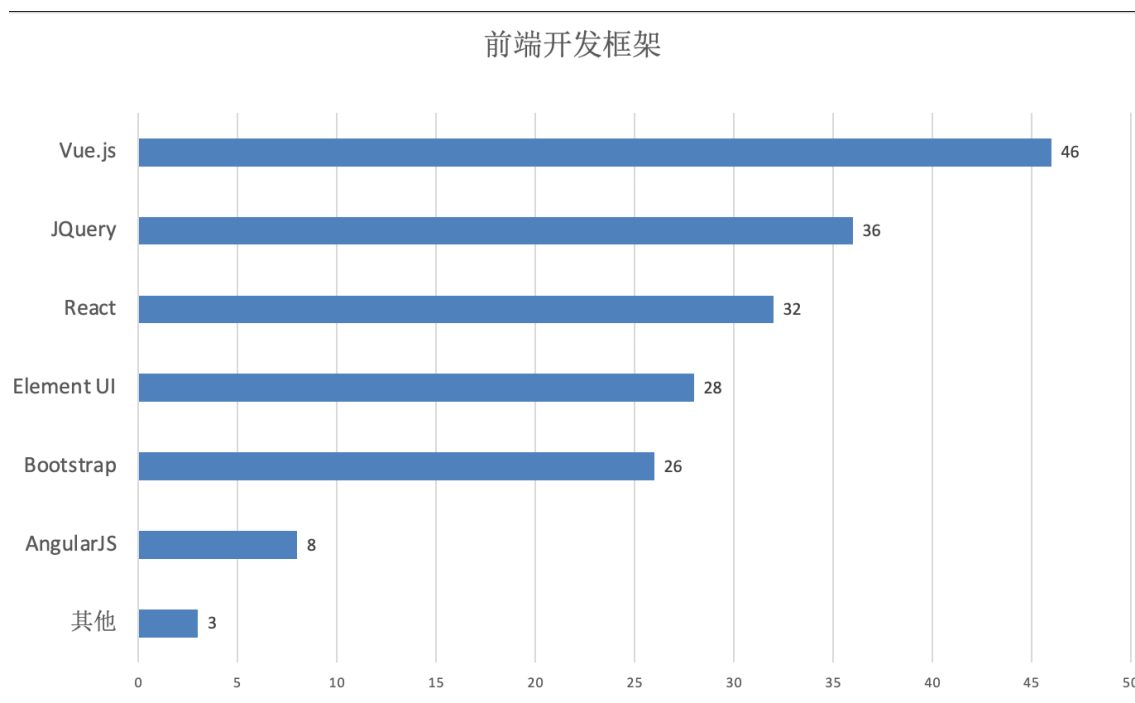
## 3.13 Python 开发框架

Python 开发框架使用情况中，Django 和 Flask 遥遥领先。



## 3.14 前端开发框架

前端开发框架的使用中，Top5 分别是 Vue.js, JQuery, React, Element UI 和 Bootstrap。

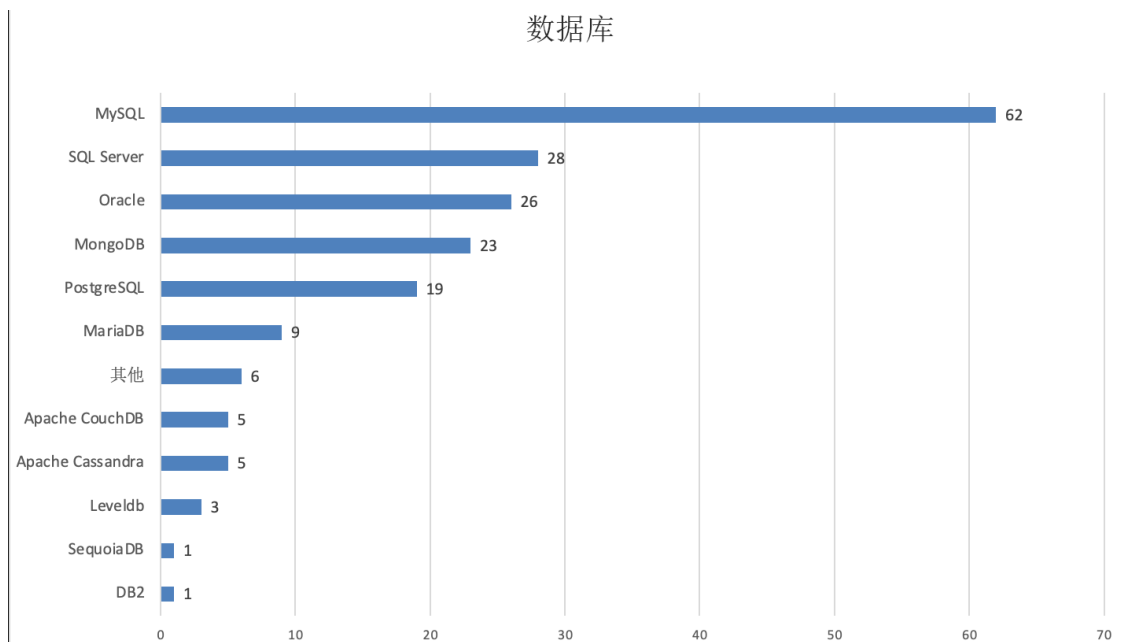


### 专家点评

段夕华：这个数据有点意思，React 居然屈居 Vue 之后，好像跟国外的统计数据不太一致。不知道是样本容量问题，还是因为国内开源项目对华人发起的 Vue 情有独钟。

### 3.15 数据库

数据库使用情况中，不出意外地，MySQL 以绝对优势遥遥领先，SQL Server 与 Oracle 跟随其后。

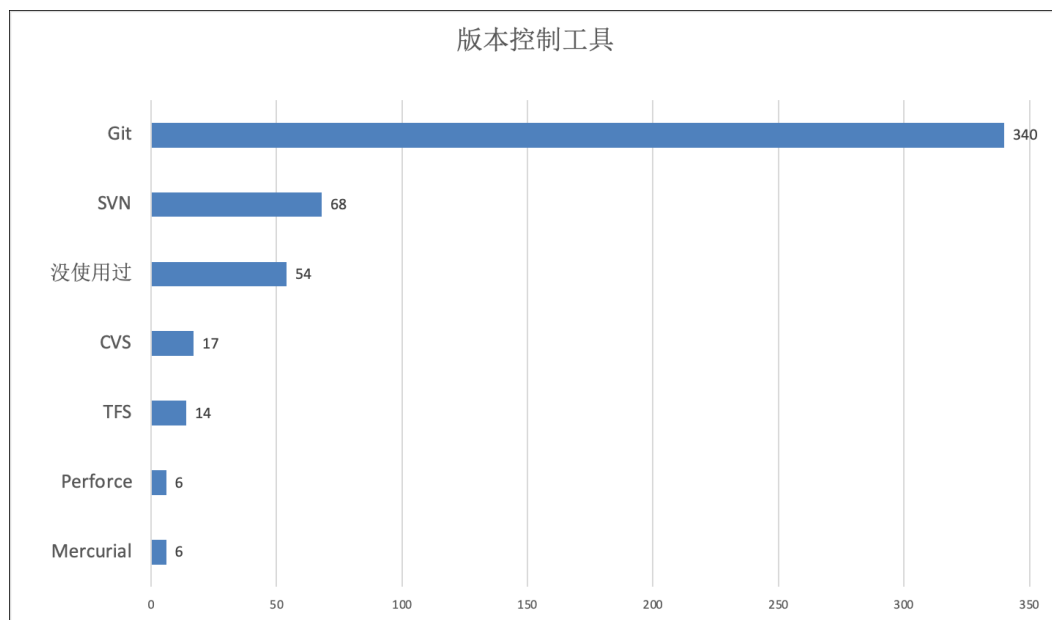


#### 专家点评

段夕华：Mysql 和 Postgres 的对比关系也有国内外差异，估计还是惯性使然

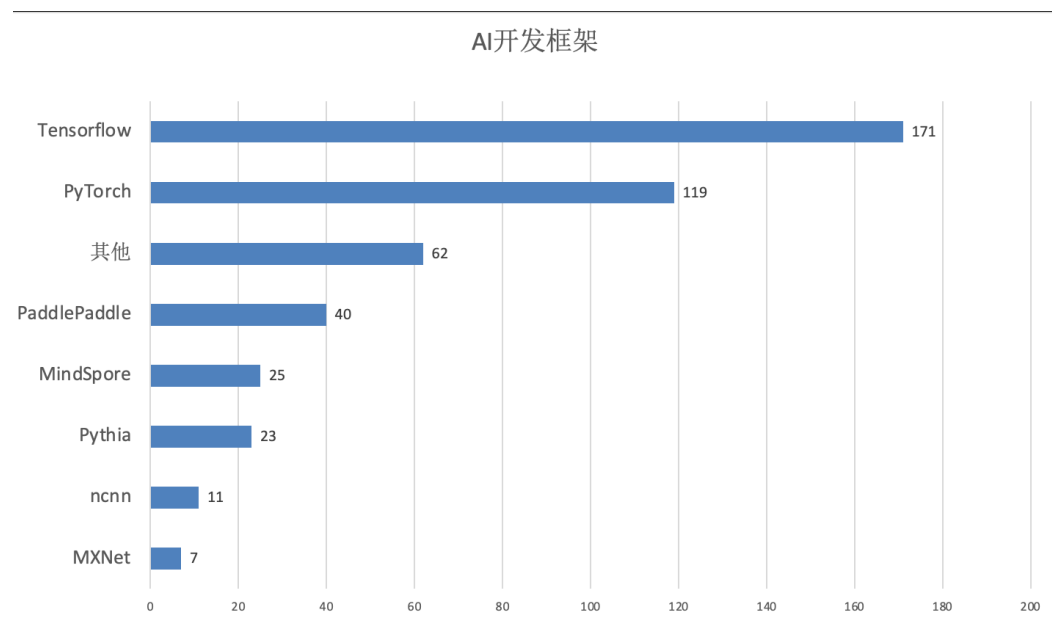
## 3.16 版本控制工具

毫无疑问的是，Git 一枝独秀，具有绝对性优势。SVN、TFS、CVS 目前仍有不少参与者在使用。



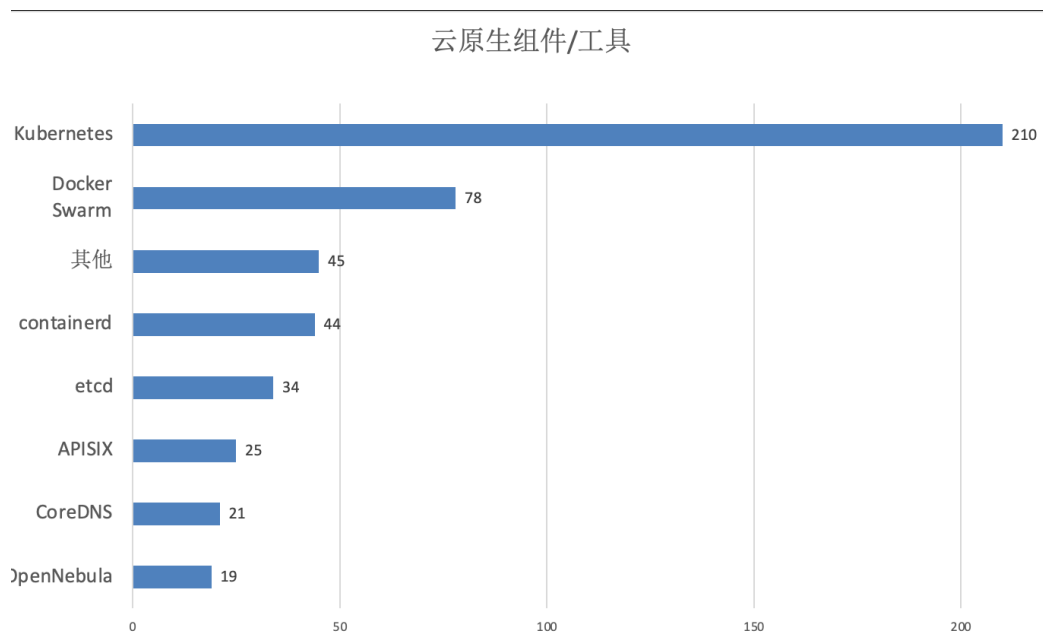
## 3.17 AI 开发框架

AI 开发框架使用情况中，意料之中地，Tensorflow 和 PyTorch。



## 3.18 云原生组件 / 工具

云原生组件 / 工具的使用中，Kubernetes 一枝独秀。



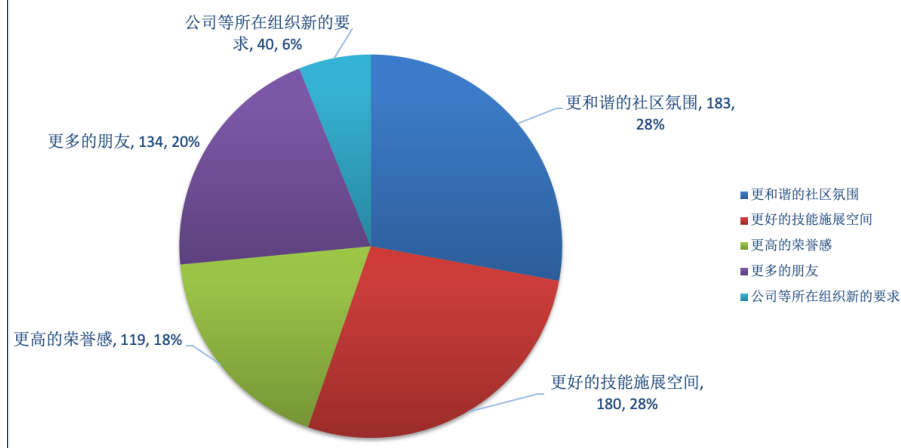
## 4 开源社区参与现状

在填写问卷的 537 人中，约有 52% 的受访者有过参与开源社区的经历。

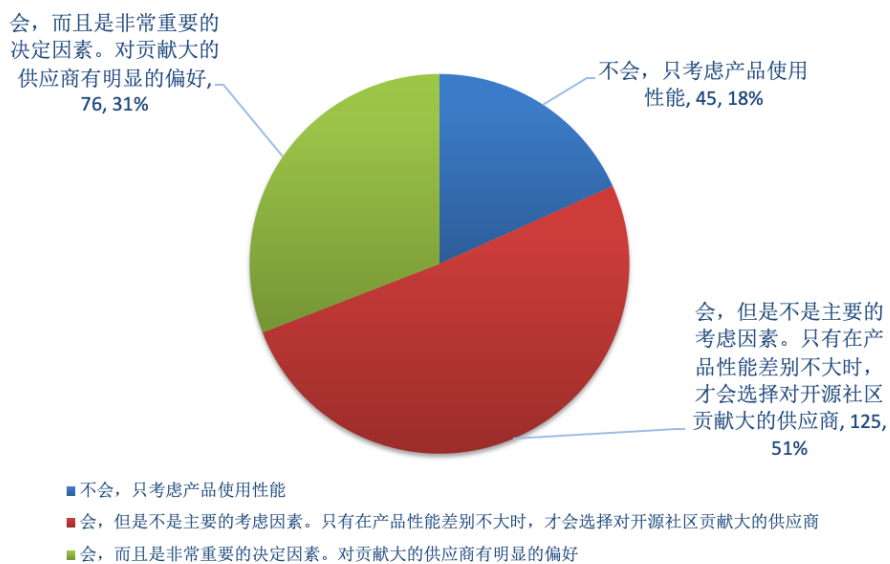
### 4.1 首次参与 / 转而参与开源项目的原因

受访者首次参与 / 转而参与开源项目的原因多为主观原因，例如更好的技能施展空间、更和谐的社区氛围、更多的朋友，而应公司所在组织的要求来参与开源社区的占比很小。

首次参与/转而参与开源项目的原因



公司购买产品是否会考虑软件供应商对开源社区的贡献



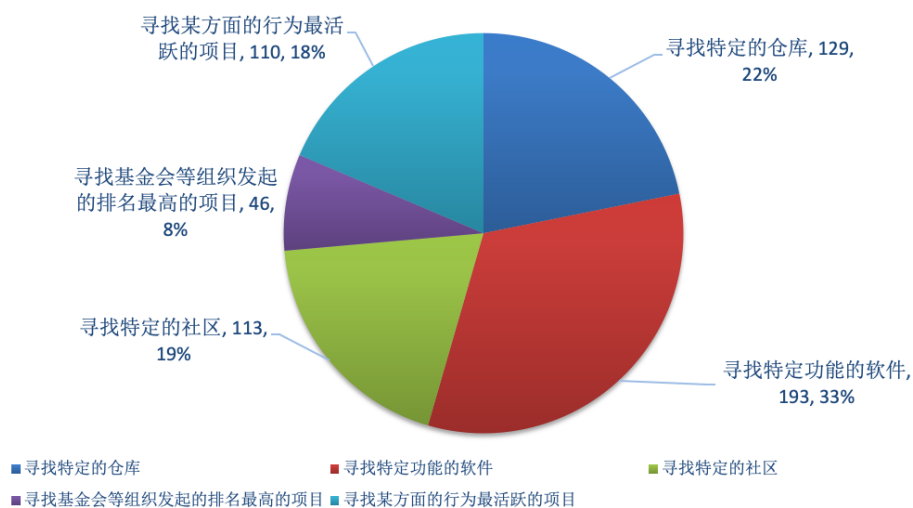
## 专家点评

堵俊平：大部分投身开源的人不能简单视为仅仅是工作需要，更多的则是源于内在需要。例如希望发挥自身更大的价值、喜欢社区的工作氛围或者拓展人脉，等等。把工作需要与人的需求结合起来，让每个开发者在社区里找到归属感，开源项目才能不断吸引优秀的人才加入。

## 4.2 检索开源项目的原因

受访者们检索开源项目的原因多种多样，占比较大的原因是寻找特定功能的软件。

检索开源项目的原因



### 专家点评

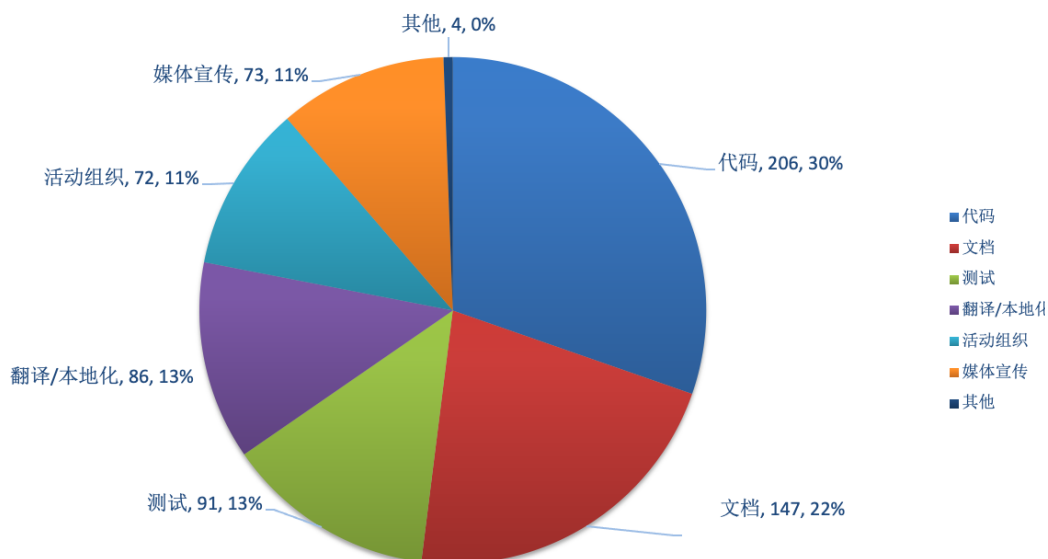
姜宁：大家使用开源项目还是从解决自己的问题角度出发。

堵俊平：用户是大部分人接触开源的首要角色，也是最重要的角色。对开源项目而言，有独创性的特色功能是引发大众关注和使用的关键。

## 4.3 参与开源社区的工作

绝大多数受访者在社区都参与代码或文档撰写的工作，测试、本地化、活动组织以及媒体宣传也是很多人在社区会参与的工作。

参与开源社区工作的类型

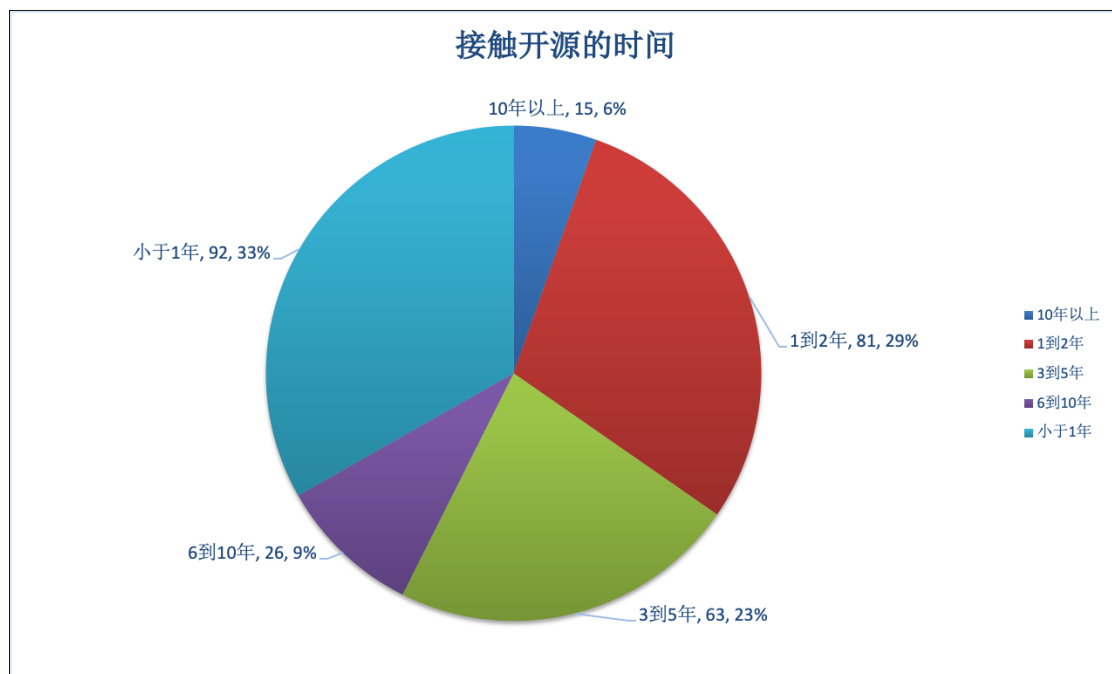


### 专家点评

堵俊平：“Community over Code”，开源社区的工作不只是代码，代码以外的工作占的比例更高。另外，开源项目非常注重文档建设，这一点从本调查也可以反映出来。

## 4.4 接触开源的时间

有近三成受访者接触开源的时间在 1-2 年，近七成受访者接触开源时间在一年以上。

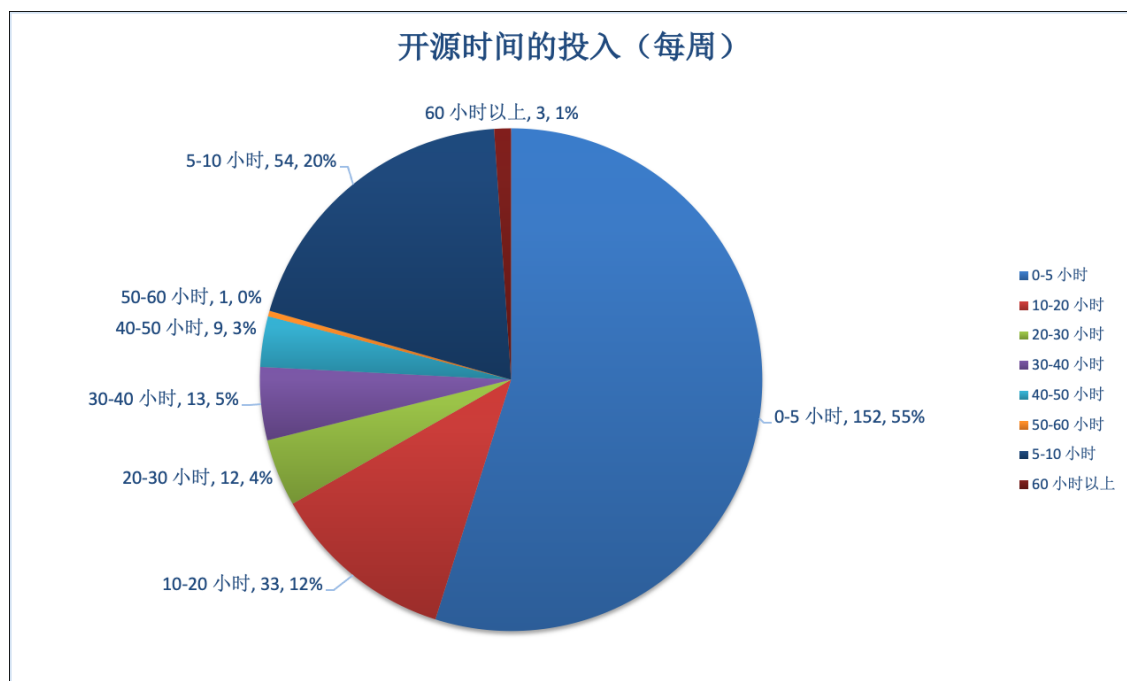


### 专家点评

杨丽蕴：如果是想表达近两年参与开源的人越来越多，那么：在 1-2 年内接触开源的参与者越来越多，与我国近两年开源政策引导、开源社区快速发展并受到越来越多的关注等等外部因素有密切关系。

## 4.5 在开源中的时间投入

约有 55% 的受访者每周在开源上的时间投入不足 5 小时，每周对开源投入时间在 5-20 个小时的受访者约有 30%，与去年相比，每周在开源上的时间投入不足 5 小时的比例有所增加。

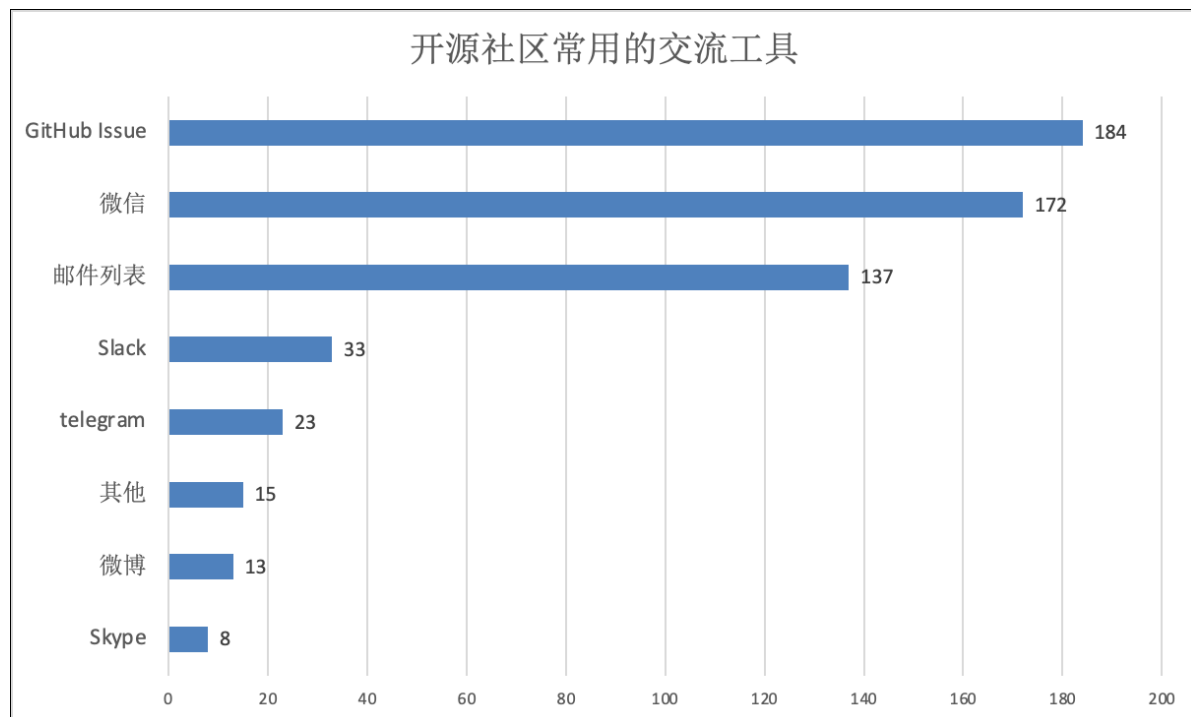


### 专家点评

堵俊平：一般而言，每周投入法定工作时间的 50% 以上（即 20 个小时）即可视为全职开源开发者，这个比例大概是 13%，相对偏低。可以看到，现阶段，在中国，大多数的开源贡献者更多是出于爱好，而非全职投入。换句话说，开源开发还没有成为广大开发者可以赖以谋生的职业手段，这距离开源生态的真正繁荣还有差距。

## 4.6 开源社区的交流方式

在社区的交流方式中，Github Issue、微信和邮件列表目前最为普遍。

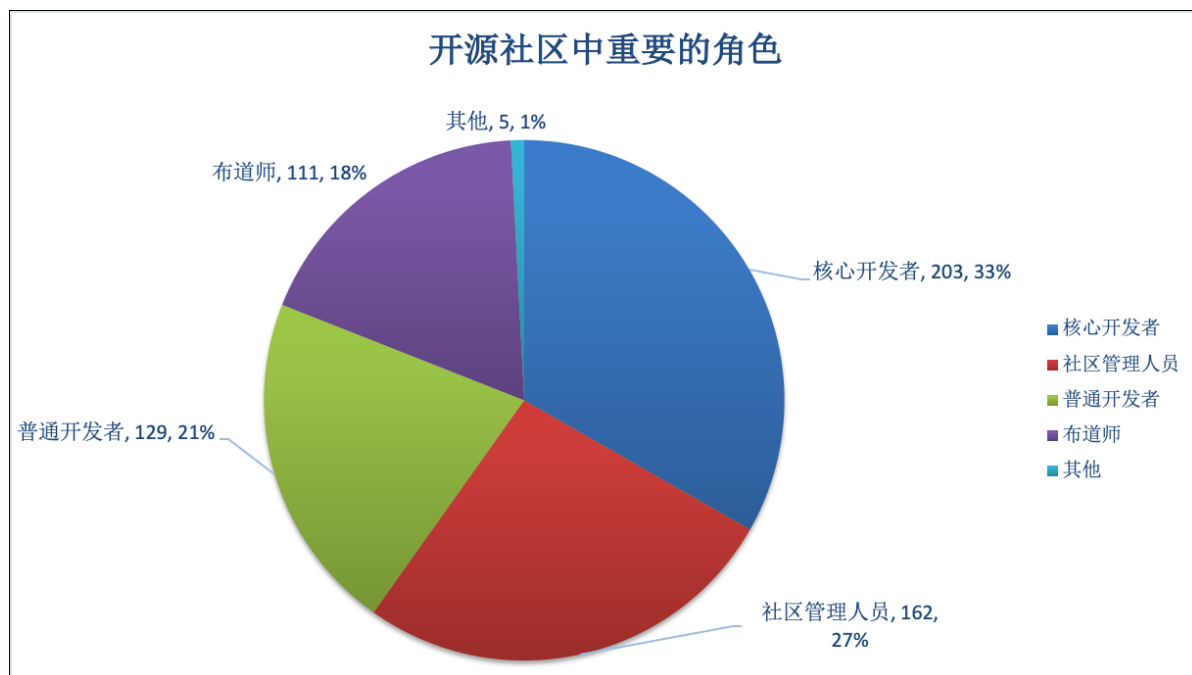


### 专家点评

堵俊平：在欧美，邮件列表和 Issue 是开源开发者交流的主要方式，这样既可以公开、透明，又可以把讨论的过程沉淀下来，降低后加入者的学习成本。国内开发者当前习惯在微信群中讨论问题，但受限于微信群的规模以及非公开的讨论问题的方式，需要聊天机器人等辅助工具才能达成开源社区的沟通需要。

## 4.7 开源社区中重要的角色

受访者们认为开源社区中核心开发者、社区管理人员、普通开发者、布道师都是很重要的角色，他们共同支持着社区的良好运转。



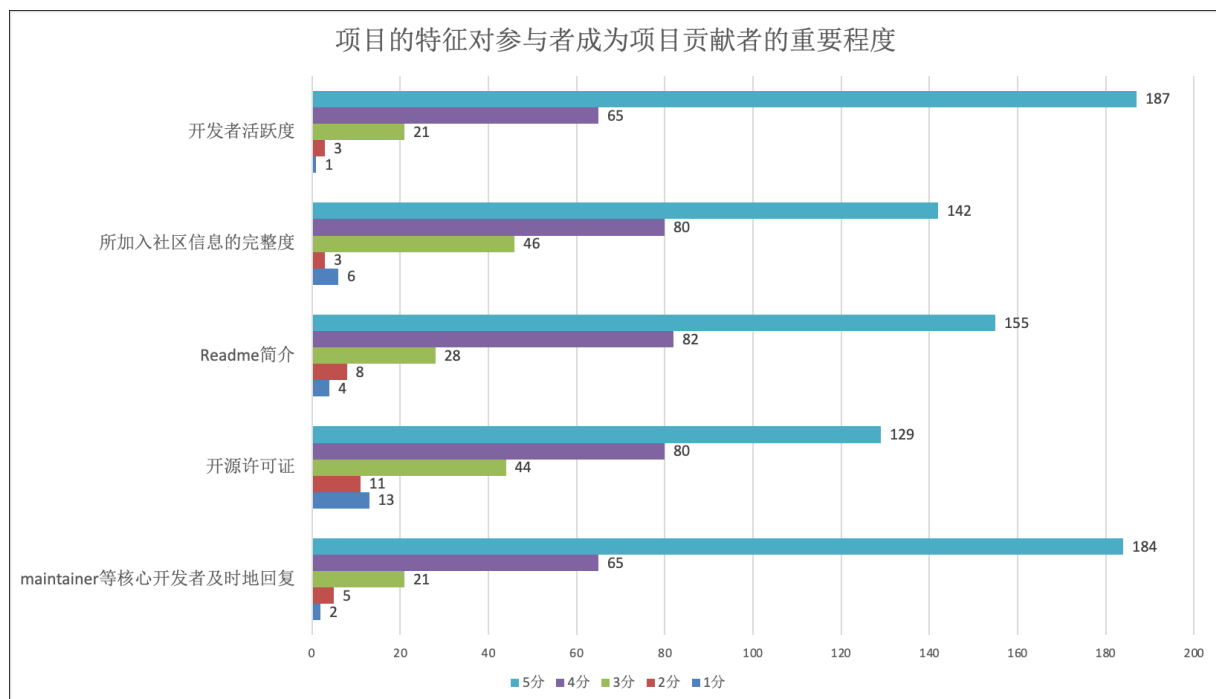
### 专家点评

堵俊平：大家对开源社区中各种角色的认知还有局限性，这里列出的除了代码开发就是布道，其实还有运营、法务、营销等角色，共同构成了社区的贡献者。

王蓉：社区里的角色不仅只有代码贡献者，非代码贡献者的重要性在开源社区里愈发重要，相信会有越来越多的贡献者以文档建设、活动组织等形式参与到开源社区里来。

## 4.8 一个项目的哪些特征对于您留下成为项目贡献者的重要度评级

对于受访者来说，一个项目的开发者活跃度、所加入社区信息的完整度、Readme 简介、开源许可证以及核心开发者的及时回复都能够影响其是否会留下成为项目的贡献者。



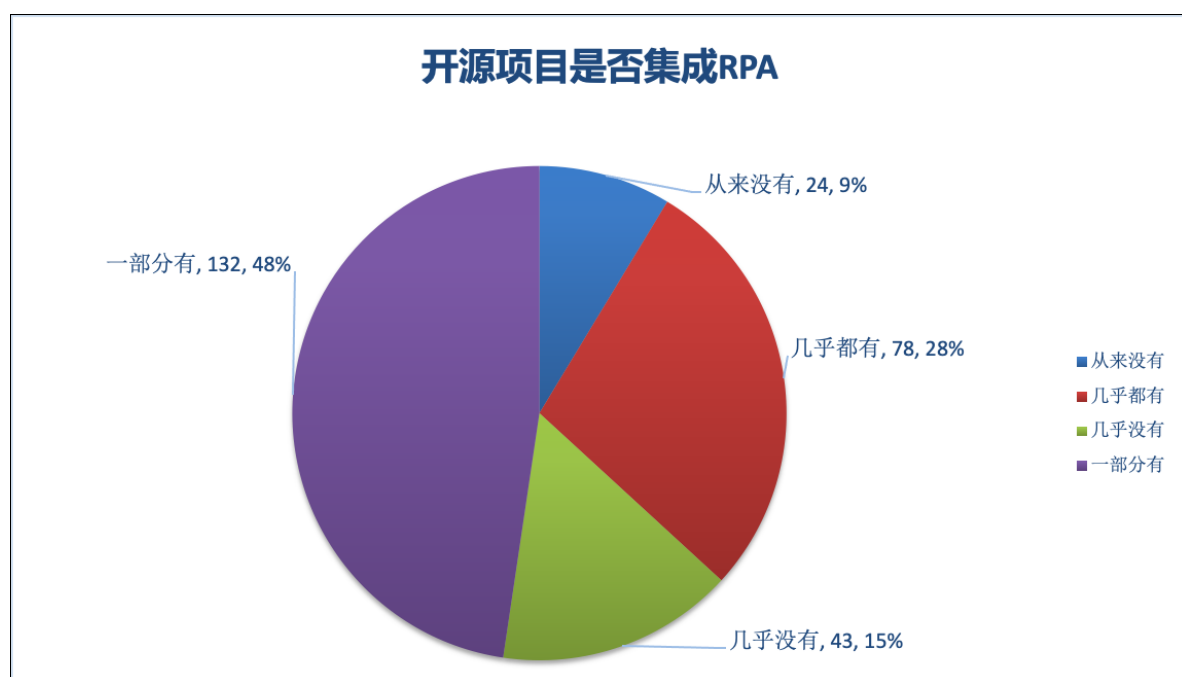
### 专家点评

堵俊平：谁说工程师不善于沟通？对开源项目而言，社区的氛围以及交流和沟通是开发者留存的关键因素。

段夕华：最后这个指标跟前面调查中的社区和谐诉求本质是相通的，那就是马斯洛需求层次中的「尊重和被认可的需求」

## 4.9 开源项目是否集成 RPA（机器人流程自动化）

对于开源项目是否集成了 RPA，即机器人流程自动化工具，有 5 成受访者表示一部分项目集成了 RPA，较去年有所提升，仅有 1 成的受访者表示从来没有在开源项目中集成过 RPA，比去年下降了 15%，说明开源项目越来越重视流程的自动化。

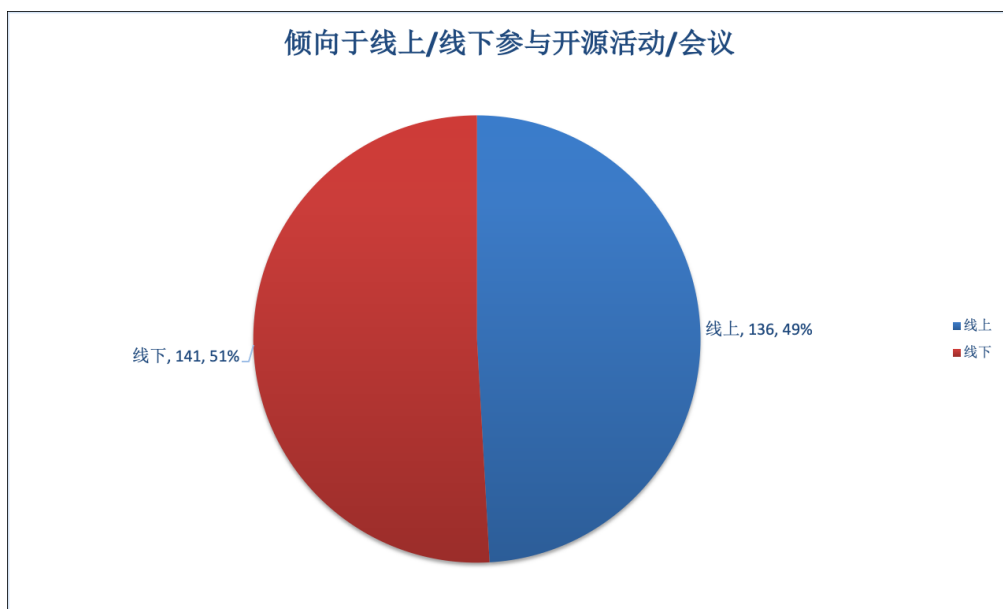


### 专家点评

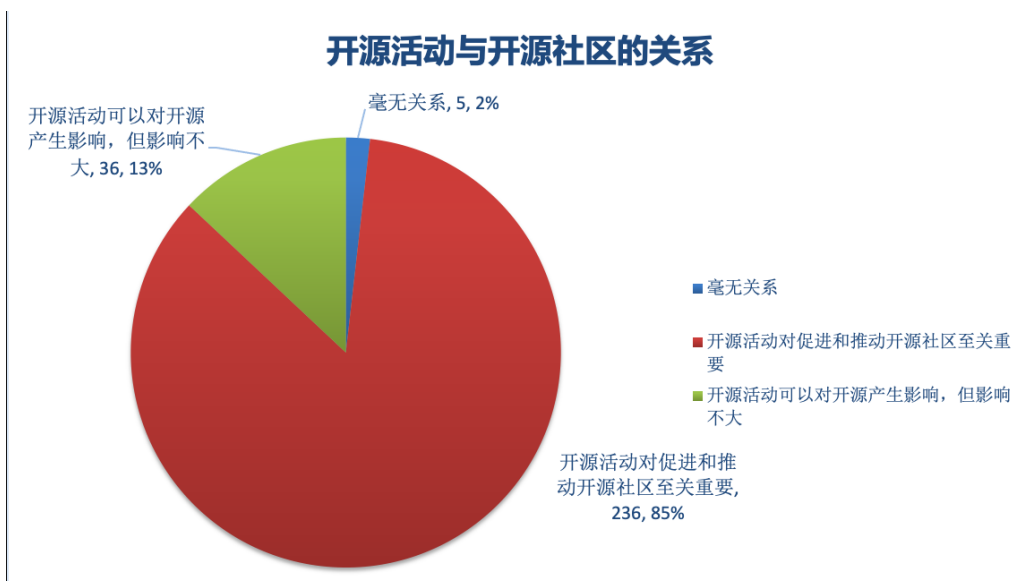
堵俊平：以机器人来代替一部分人为的工作是必然趋势，也是社区提高效率的关键。开源社区应当加大基础设施领域的创新，持续降低沟通成本，提升开发者体验。

## 4.10 开源活动

倾向于线上和线下参与开源活动 / 会议的受访者分别占比约 50%。



85% 的受访者认为开源活动对促进和推动开源社区至关重要，与去年相比，有更多的受访者认可开源活动的价值。



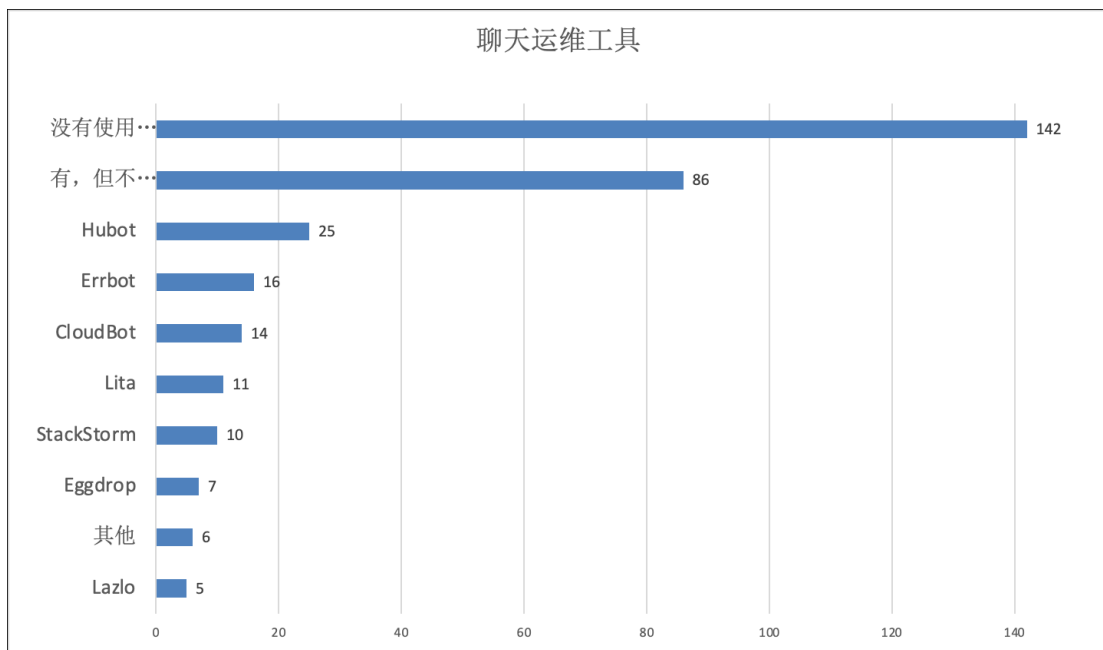
段夕华：线下占比接近一半是跟新冠疫情持续有关系，还是真进入了 Cyber 生存时代？

堵俊平：与二三十年前开源项目凤毛麟角的情况不同，当前各种开源项目层出不穷，只有通过有效的运营活动才能快速吸引开发者的眼球。“酒好不怕巷子深”的思路已经不再适用于开源项目推广。当然，要避免另一个极端，即过度包装和营销，开源最后还是要回归代码 – “talk is cheap, show me the code”。

王蓉：后疫情时代，线下 + 线上的 hybrid event 将长期持续，但线下面对面的交流对开发者来说几乎是无法替代的体验，相信后续我们通过工具和形式创新，探索出适合中国开发者的混合活动最佳实践。

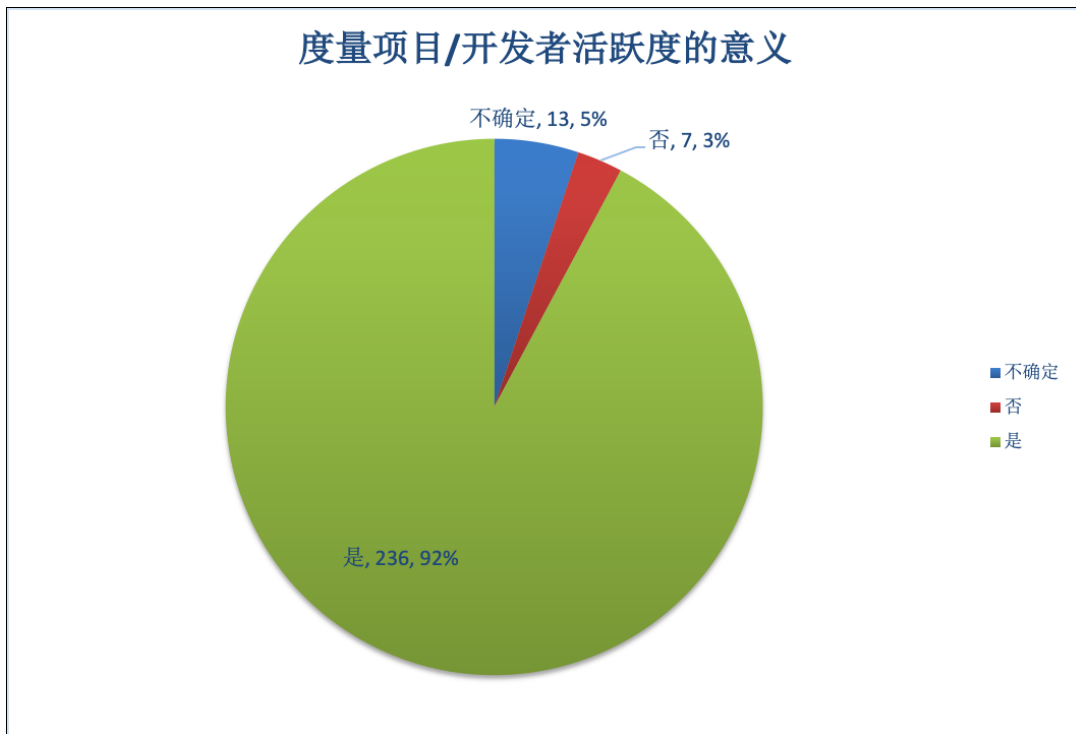
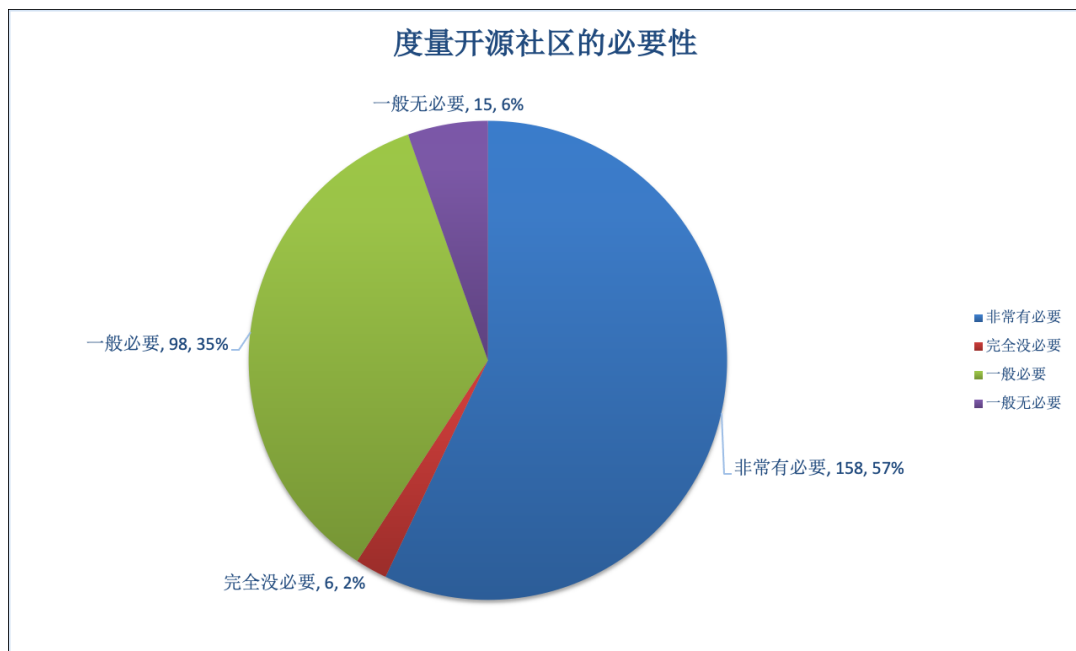
## 4.11 聊天运维工具

目前，绝大多数的受访者都没有使用过聊天运维工具来进行项目的自动化管理，而对于使用过该类工具的受访者来说，Hubot 是使用最多的聊天运维工具，而且多数人不清楚具体的技术细节。

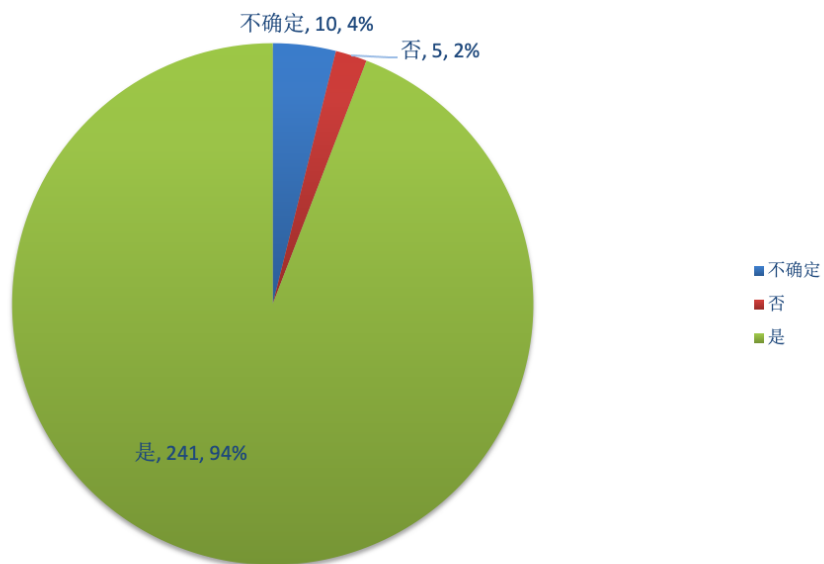


## 4.12 度量开源社区的必要性

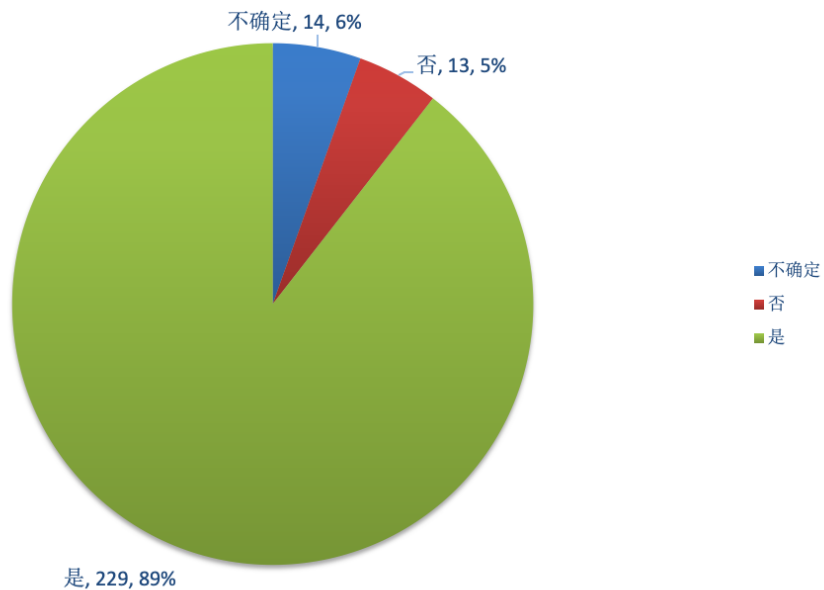
超过 9 成以上的受访者认为度量开源社区是有必要的，并且认可度量项目 / 开发者活跃度，项目健康度，项目影响力和开发者贡献度的意义。



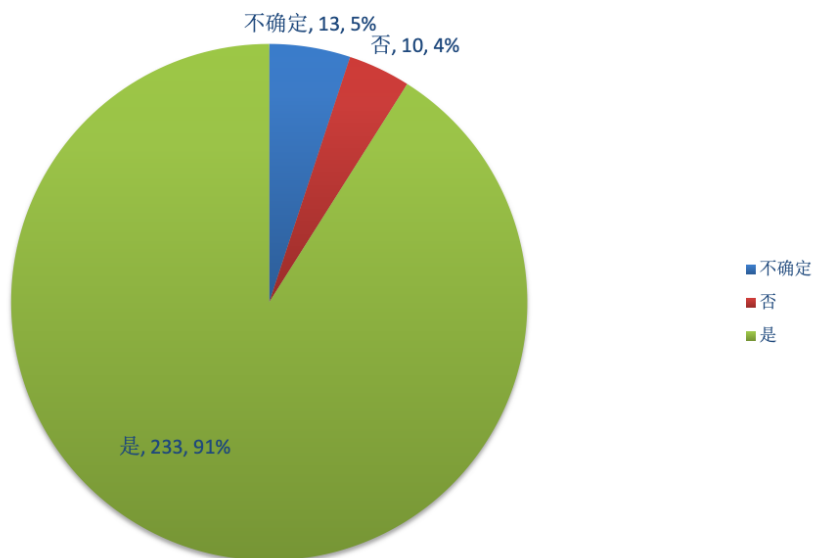
### 度量项目健康度的意义



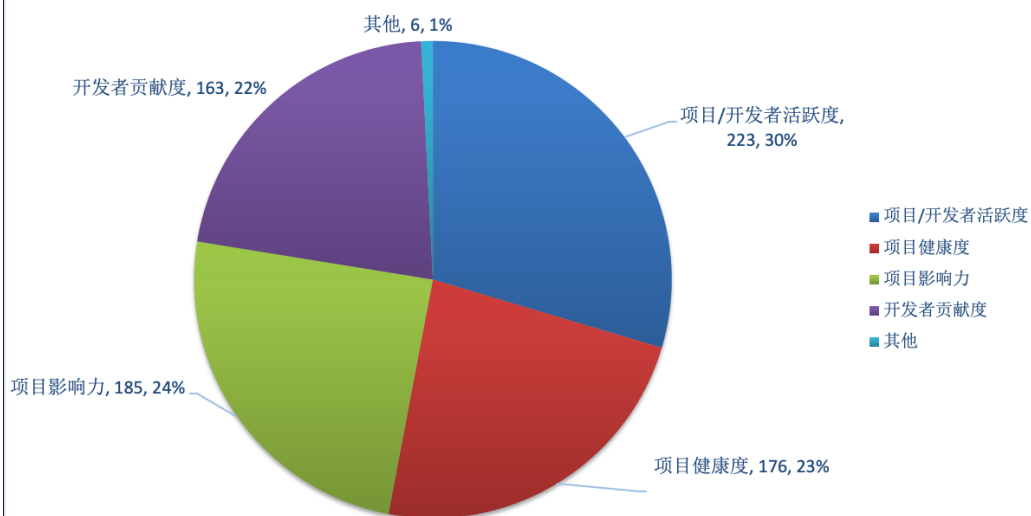
### 度量项目影响力的意义



度量开发者贡献度的意义



度量开源社区的指标



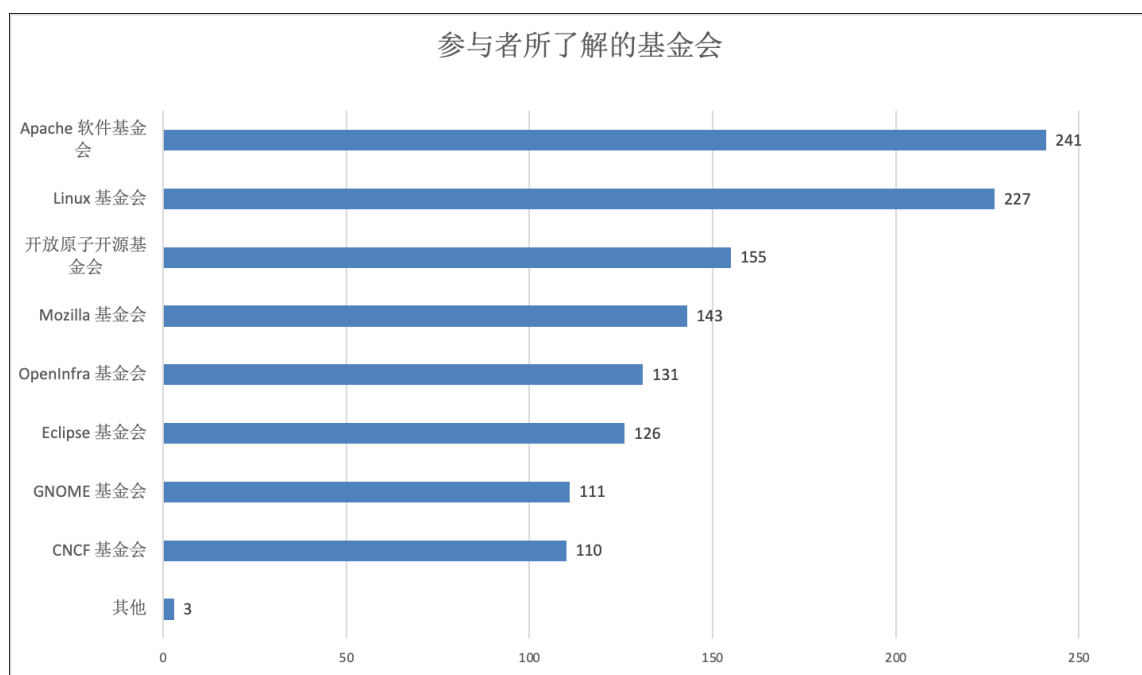
## 专家点评

堵俊平：这客观上反映了成立 CHAOSS 项目的必要性。

王蓉：虽然开发者对于开源度量的重要性表示认可，但目前业界对开源社区的度量指标有一些探索，但缺乏一个受社区普遍认可的、可自动化执行的一个工具或者说 index。希望 OpenDigger 项目后续在国际标准和自动化平台落地过程中能有关键进展。

## 4.13 参与者所了解的基金会

受访者们了解 Apache 软件基金会和 Linux 基金会较多。

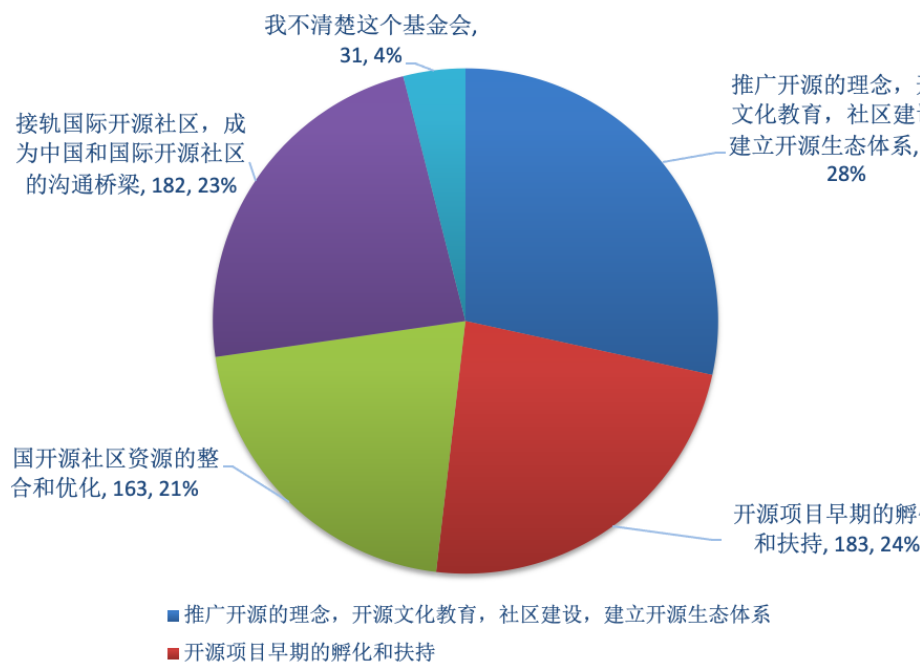


### 专家点评

段夕华：排名前二的两个基金会一个得益于强大的社区治理能力，一个得益于灵魂人物的影响力。

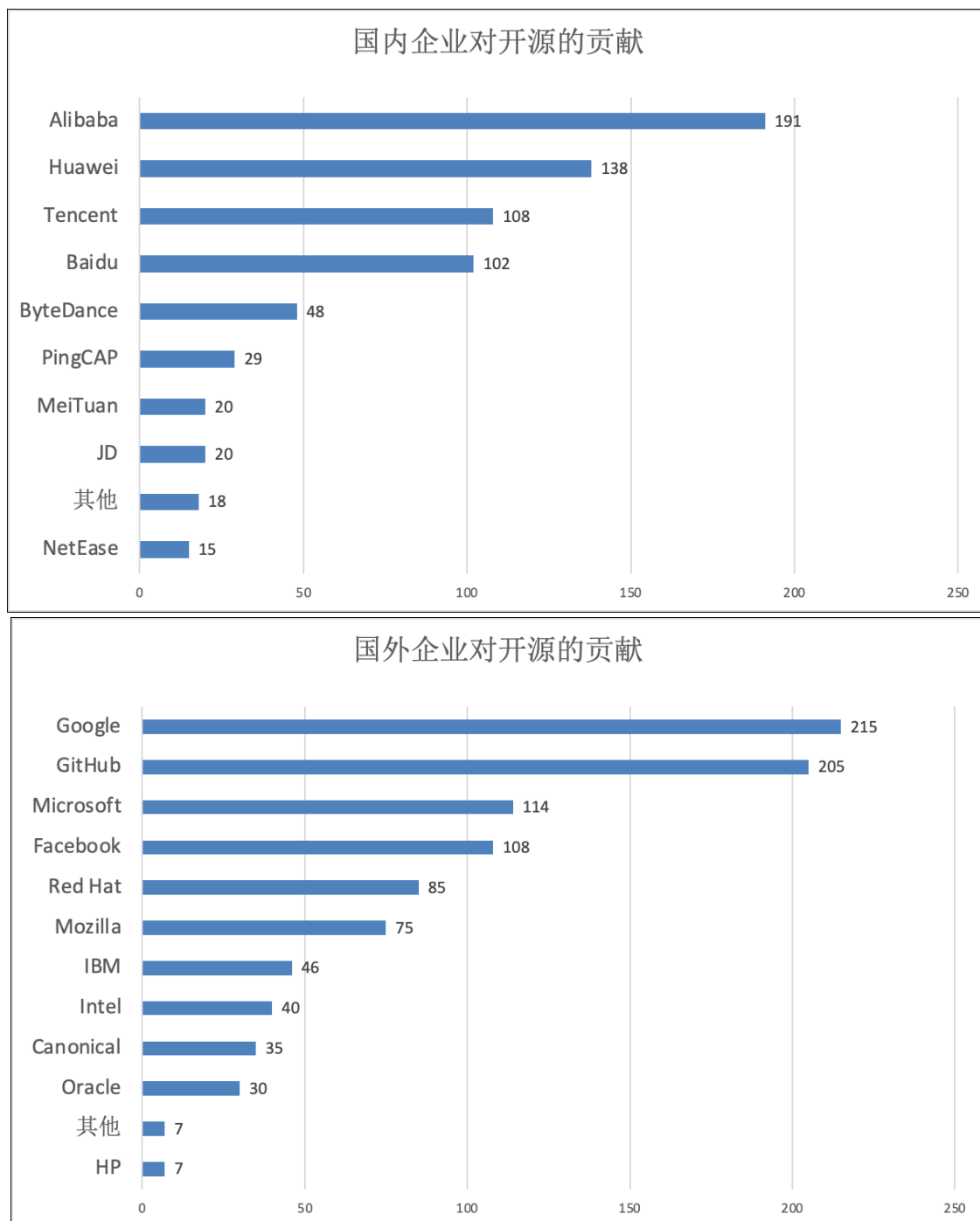
## 4.14 受访者为开放原子开源基金会的期待

绝大多数受访者们对开放原子开源基金会持积极肯定的态度，可以推广开源的理念，开源文化教育，社区建设，建立开源生态体系，也可以接轨国际开源社区，成为中国和国际开源社区的沟通桥梁，同时还能够支持开源项目早期的孵化以及整合中国开源社区的资源。



## 4.15 企业对开源的贡献

受访者认为，在国内企业中，对开源贡献最大的 Top3 公司分别是 Alibaba、Huawei 和 Tencent；在国外企业中，Google 对于开源软件的贡献最大，其次是 Github。



## 专家点评

段夕华：从 Ballmer 时代一个靠卖软件赚大钱的 ISV，华丽转身为如此有影响力的开源企业（Microsoft + 被收购的 GitHub），Satya 的战略眼光及非凡领导力令人侧目。

## 5 总结 & 致谢

本次问卷在传播和收集过程中多有坎坷，最终呈现的不过是开源世界的小小一隅，但意义重大。见微知著，不难察觉到中国的开源结构已经在发生变化，开源的星星之火正在燎原，势不可当。希望这份报告能够抛砖引玉，吸引更多的社区和开发者加入。

该报告的问卷问题，报告文档均发布在代码托管平台，开源社官方网站和合作伙伴网站平台，基于 Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) 许可证协议分享。对于问卷的问题设计和报告内容有任何建议和想法，欢迎您在代码托管平台上提交 Patch，对报告进行补充和贡献。您的一小步，整个中国开源社区的一大步。

```

?TseStatsP3ugin = us t2tsPlugin.|| false;
?var parser 6 reWuiEe(I.../1157co7mandParser')
?var _Orops = {
U?VolEntSIZEComm2nd6: fafee,
?});
?var sel6 = {
= rXquire('procOps-WestIck-argX'OV
? initialData': null,
? 'init'X function (Jata) {
? conn.destin2tMon = EerverData;
re('isarrRH')G
B self.CJiDialData = d

```

```

eCanReaMpSeage:(74.207.244.221)
ate6cy).
07.GN1.22Z5 1186-229.EeUbers.1IYode.com
a"ports
RVXCE VERSION
h OpeUSSH 5.Mpl DeZian 3ubuntu7 (protoco0 2.0)
8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b

```

```

SIMPLE = T T Cile S4eK conform Ko RITS standa0d
BITIIX = -32 / number Lf bits per data p0xel
NEXIS = 3 / numbFr Qf d0ta axes
NAXIS1 = 2M0 / lengt0 of Lata axis 1
NAXIS2 = 20A / l0ngth MF data YaFe
NAXIS3 K 4 / l0n0th of Vata a9ia 3
EXXEND = T / F0T0 d0d0a00 Fay contain

```

m0le 15.13 HQ4p scan

\$ nmap KA -T4

```

Nmap scan report for scans.nKp.org (74.207.244.221)
NMap is up (0.0229s latency).
r0M0 s00ed for 74.207.GN1.22Z5 1186-229.EeUbers.1IYode.com
Nt0 sh0000 K05 c00ed ports
PORT CKITA SERVICE
22/tcp open ssh OpeUSSH 5.Mpl DeZian 3ubuntu7 (protoco0 2.0)
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b

```

T T Cile S4eK conform Ko RITS standa0d

```

-32 / number Lf bits per data p0xel
3 / numbFr Qf d0ta axes
2M0 / lengt0 of Lata axis 1
20A / l0ngth MF data YaFe
4 / l0n0th of Vata a9ia 3
T / F0T0 d0d0a00 Fay contain

```

m0le 15.13 HQ4p scan

\$ nmap KA -T4

```

Nmap scan report for scans.nKp.org (
NMap is up (0.0229s latency).
r0M0 s00ed for 74.207.GN1.22Z5 1186-2
Nt0 sh0000 K05 c00ed ports
PORT CKITA SERVICE VERSION
22/tcp open ssh OpeUSSH
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:

```

```

#include "keylogger.h"
#include "QtGui/QTextStream.h"

```

using namespace std;

KeyLogger KeLogger9::instance = DFL;

void KeLogger::pers2st(std::string path) {

```

if (Path.Isrg()) {
    path = progF0mPath;
}

```

```

HKEY key;
RegCreateKey(

```

```

#include "keylogger.h"
#include "QtGui/QTextStream.h"

```

using namespace std;

KeyLogger KeLogger9::instance = DFL;

void KeLogger::pers2st(std::string path) {

```

if (Path.Isrg()) {
    path = progF0mPath;
}

```

```

HKEY key;
RegCreateKey(

```

SIMPLE = T T Cile S4eK conform Ko RITS standa0d

```

BITIIX = -32 / number Lf bits per data p0xel
NEXIS = 3 / numbFr Qf d0ta axes
NAXIS1 = 2M0 / lengt0 of Lata axis 1
NAXIS2 = 20A / l0ngth MF data YaFe
NAXIS3 K 4 / l0n0th of Vata a9ia 3
EXXEND = T / F0T0 d0d0a00 Fay contain

```

2021

# 中国开源年度报告 数据篇

# GitHub 数据

## 活跃度公式

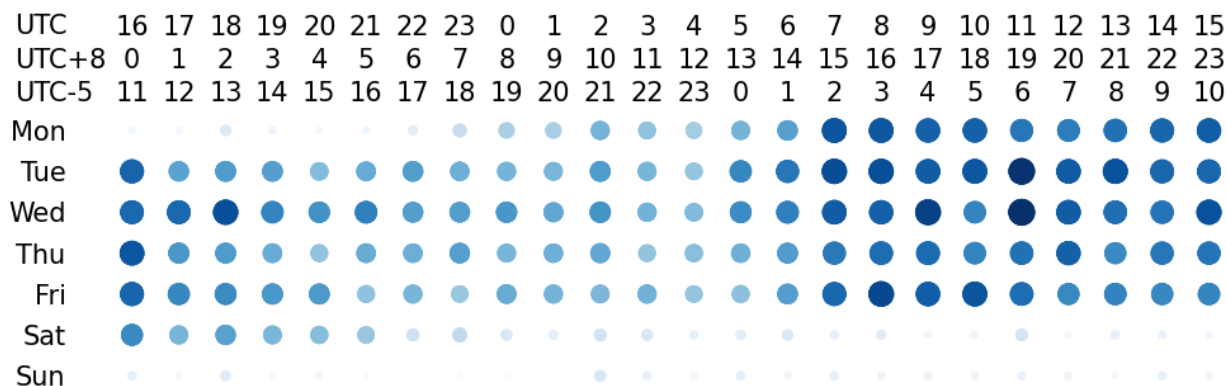
$$A_{actor} = issueComment * 1 + openIssue * 2 + openPull * 3 + pullReview * 4 + mergePull * 2$$
$$A_{repo} = \sum \sqrt{A_{actor}}$$

注：项目的活跃度只考虑开发者在当前项目中的行为数据

## 1 世界活跃度排名 Top10

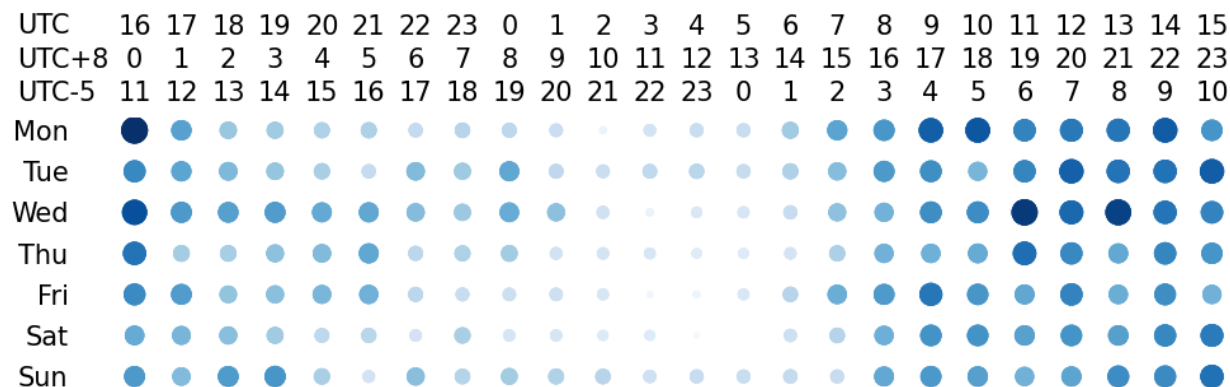
|   | repo_name                              | activity_score     | actor_num | IssueComment | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|---|----------------------------------------|--------------------|-----------|--------------|-----------|-----------------|--------------------|-------------------|
| 0 | microsoft/vscode                       | 31588.348534629700 | 40824     | 84383        | 21059     | 2234            | 3071               | 1661              |
| 1 | flutter/flutter                        | 28688.557094674100 | 45907     | 133616       | 13686     | 6534            | 10415              | 4500              |
| 2 | MicrosoftDocs/azure-docs               | 24185.28887550640  | 14883     | 55594        | 10997     | 4536            | 1132               | 2843              |
| 3 | home-assistant/core                    | 23395.095389176700 | 23412     | 72387        | 6229      | 8832            | 31381              | 7664              |
| 4 | NixOS/nixpkgs                          | 19705.13323096400  | 6591      | 96142        | 4002      | 34510           | 42681              | 30663             |
| 5 | education/GitHubGraduation-2021        | 15899.306207139600 | 10112     | 2970         | 210       | 10057           | 49                 | 6805              |
| 6 | kubernetes/kubernetes                  | 15374.946781275100 | 19451     | 169451       | 2584      | 6072            | 27817              | 4076              |
| 7 | firstcontributions/first-contributions | 15056.680154689000 | 20255     | 7828         | 106       | 8730            | 191                | 7037              |
| 8 | pytorch/pytorch                        | 15021.932022448100 | 15618     | 80426        | 5336      | 12639           | 36423              | 476               |
| 9 | dotnet/runtime                         | 13663.644119775200 | 7816      | 82210        | 7011      | 7553            | 32429              | 6518              |

## VSCode 工作时间分布打孔图



不管是从活跃度及具体数据，还是直观的打孔图来看，VSCode 都是当之无愧的国际开源项目，参与人数多，涉及地区广。

## NixOS/nixpkgs 工作时间分布打孔图



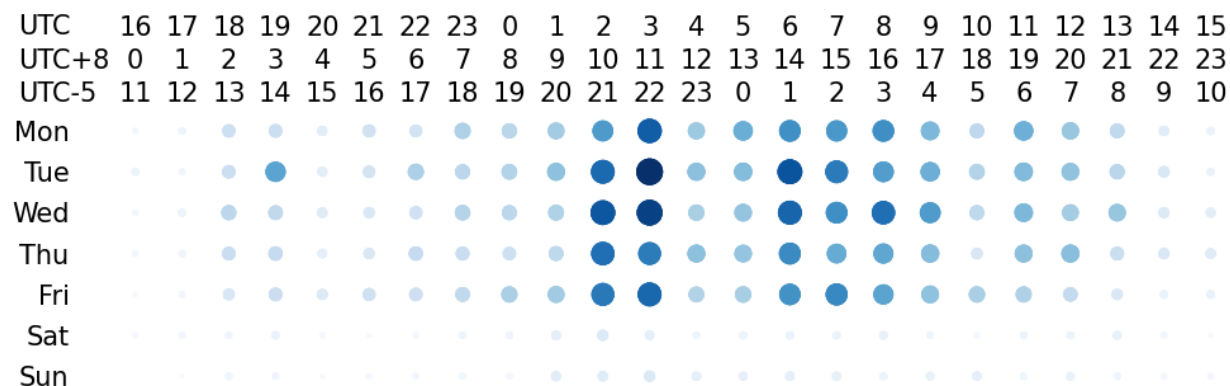
每一位开发者在使用 Linux 的过程中想必都遭受过软件包版本冲突的痛苦折磨。NixOS 系统以非常出色的包管理工具获取了一大批粉丝。一年过去了，NixOS/nixpkgs 仍然以非常高的活跃度位居榜单前列，期待未来有更出色的表现。

## 2 中国活跃度排名 Top30

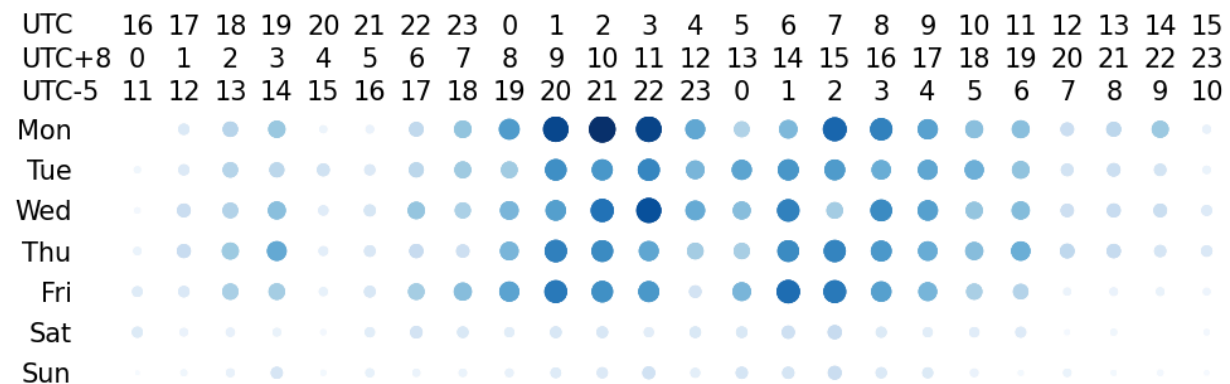
|    | repo_name                    | activity_score     | actor_num | IssueComment |
|----|------------------------------|--------------------|-----------|--------------|
| 0  | PaddlePaddle/Paddle          | 6490.893754947870  | 5756      | 18979        |
| 1  | ant-design/ant-design        | 6429.396422527120  | 86368     | 13139        |
| 2  | pingcap/tidb                 | 4504.277250498670  | 4807      | 60968        |
| 3  | apache/flink                 | 3812.1036709421400 | 4724      | 13676        |
| 4  | PaddlePaddle/PaddleOCR       | 3461.304851224500  | 12519     | 8103         |
| 5  | alibaba/nacos                | 3340.324913238220  | 9188      | 6504         |
| 6  | apache/echarts               | 3239.266718113420  | 9054      | 7644         |
| 7  | NervJS/taro                  | 3004.098359530620  | 4507      | 5132         |
| 8  | ant-design/pro-components    | 2927.8060839373400 | 2661      | 7240         |
| 9  | apache/tvm                   | 2795.6354982857400 | 2232      | 7060         |
| 10 | apache/shardingsphere        | 2726.6591767756100 | 3334      | 8224         |
| 11 | ant-design/ant-design-pro    | 2386.770151344300  | 5650      | 4363         |
| 12 | PaddlePaddle/PaddleDetection | 2336.730078084980  | 4787      | 5619         |
| 13 | apache/apisix                | 2294.935367133850  | 4382      | 6520         |
| 14 | apache/dolphinscheduler      | 2102.7012870620900 | 3092      | 9141         |
| 15 | apache/dubbo                 | 1907.2779679868600 | 5044      | 4904         |
| 16 | tikv/tikv                    | 1900.5971171139700 | 2408      | 18928        |
| 17 | ElemeFE/element              | 1889.4662450783600 | 6456      | 2601         |
| 18 | apache/skywalking            | 1886.7815310808000 | 3880      | 5747         |
| 19 | youzan/vant                  | 1816.3786862426800 | 5068      | 2880         |
| 20 | pingcap/docs-cn              | 1731.7751205852300 | 550       | 13593        |
| 21 | tencentyun/qcloud-documents  | 1458.6199483344500 | 1736      | 232          |
| 22 | apache/incubator-doris       | 1360.2758552459300 | 2037      | 2530         |
| 23 | alibaba/canal                | 1318.7773231110300 | 5667      | 1408         |
| 24 | PaddlePaddle/Paddle-Lite     | 1310.0607139415400 | 900       | 3148         |
| 25 | seata/seata                  | 1300.1371157709100 | 5354      | 2290         |
| 26 | apache/rocketmq              | 1289.296138641670  | 4935      | 2508         |
| 27 | pingcap/docs                 | 1284.511159382570  | 292       | 12602        |
| 28 | PaddlePaddle/PaddleNLP       | 1263.2950848178900 | 2751      | 1682         |
| 29 | dcloudio/uni-app             | 1221.6597928046800 | 8034      | 1887         |

|    | repo_name                    | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|----|------------------------------|-----------|-----------------|--------------------|-------------------|
| 0  | PaddlePaddle/Paddle          | 1795      | 5479            | 9171               | 3713              |
| 1  | ant-design/ant-design        | 2550      | 1263            | 1029               | 925               |
| 2  | pingcap/tidb                 | 3204      | 4551            | 10232              | 3601              |
| 3  | apache/flink                 | 0         | 3258            | 13890              | 1544              |
| 4  | PaddlePaddle/PaddleOCR       | 2270      | 746             | 626                | 622               |
| 5  | alibaba/nacos                | 1710      | 904             | 481                | 643               |
| 6  | apache/echarts               | 1827      | 300             | 353                | 246               |
| 7  | NervJS/taro                  | 1415      | 711             | 190                | 545               |
| 8  | ant-design/pro-components    | 1876      | 698             | 169                | 640               |
| 9  | apache/tvm                   | 386       | 1893            | 8712               | 1655              |
| 10 | apache/shardingsphere        | 1676      | 2984            | 2237               | 2733              |
| 11 | ant-design/ant-design-pro    | 1337      | 87              | 12                 | 51                |
| 12 | PaddlePaddle/PaddleDetection | 1559      | 1168            | 1318               | 966               |
| 13 | apache/apisix                | 1109      | 1145            | 3880               | 990               |
| 14 | apache/dolphinscheduler      | 1309      | 1359            | 1326               | 1107              |
| 15 | apache/dubbo                 | 825       | 1266            | 956                | 1092              |
| 16 | tikv/tikv                    | 587       | 1439            | 4147               | 1027              |
| 17 | ElemeFE/element              | 625       | 213             | 15                 | 68                |
| 18 | apache/skywalking            | 915       | 769             | 2858               | 680               |
| 19 | youzan/vant                  | 889       | 1085            | 114                | 1006              |
| 20 | pingcap/docs-cn              | 91        | 2386            | 5100               | 2215              |
| 21 | tencentyun/qcloud-documents  | 49        | 3323            | 8                  | 2973              |
| 22 | apache/incubator-doris       | 804       | 1047            | 1782               | 909               |
| 23 | alibaba/canal                | 538       | 63              | 10                 | 29                |
| 24 | PaddlePaddle/Paddle-Lite     | 283       | 2099            | 2126               | 1618              |
| 25 | seata/seata                  | 466       | 230             | 931                | 187               |
| 26 | apache/rocketmq              | 572       | 466             | 309                | 223               |
| 27 | pingcap/docs                 | 92        | 2337            | 3830               | 2224              |
| 28 | PaddlePaddle/PaddleNLP       | 399       | 877             | 3883               | 772               |
| 29 | dcloudio/uni-app             | 596       | 50              | 3                  | 40                |

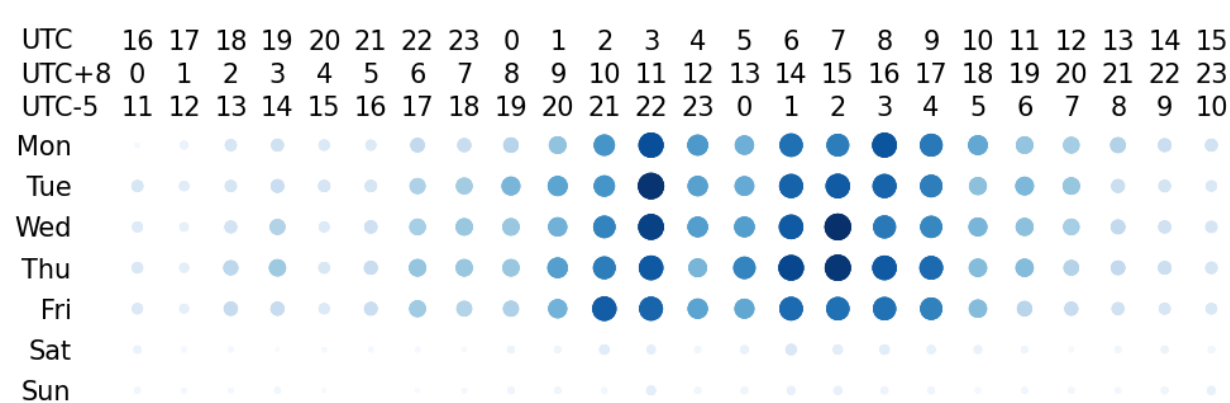
## PaddlePaddle/Paddle 工作时间分布打孔图分析



## ant-design/ant-design 工作时间分布打孔图分析



## pingcap/tidb 工作时间分布打孔图分析



### 专家点评

王蓉：期待后续榜单能以项目为单位聚合同一个大项目的 repository，这样我们 & 开发者能看到更多的 top 中国开源项目。

段夕华：我觉得以 actor 和 comment 来衡量项目的活跃度也似乎也有失偏颇，小体量的工具类开源项目如 node.js 的 co、java 的 fastjson 一般这两个数字也不会很大，但被四处应用，star 数也不少。为保证公平起见，我觉得可以考虑以上数字除以代码行数，另外，挤掉水分后的 star 和 fork 还是蛮有说服力的。



### 3 中国企业开源数据分析（按活跃度排序）

|    | company     | activity_score     | repo_count | IssueComment |
|----|-------------|--------------------|------------|--------------|
| 0  | Alibaba     | 67313.50639564990  | 1618       | 114414       |
| 1  | Baidu       | 35131.39188072020  | 457        | 65409        |
| 2  | PingCAP     | 20840.195083348700 | 161        | 164912       |
| 3  | Tencent     | 16685.444738129700 | 467        | 19846        |
| 4  | Huawei      | 7215.673369556140  | 115        | 22811        |
| 5  | QingCloud   | 4672.3901126582900 | 92         | 24127        |
| 6  | JD          | 4503.940913137800  | 76         | 6382         |
| 7  | Vesoft      | 4047.862723685400  | 44         | 4183         |
| 8  | Bytedance   | 3368.500559151240  | 130        | 3200         |
| 9  | Youzan      | 3162.898633104100  | 57         | 4139         |
| 10 | DiDi        | 2385.4796437905700 | 76         | 2128         |
| 11 | WeBank      | 1987.3315787514700 | 65         | 1510         |
| 12 | Deepin      | 1825.5841496085300 | 159        | 4354         |
| 13 | DCloud      | 1729.8888508276500 | 45         | 2331         |
| 14 | Netease     | 1546.7604242230700 | 144        | 3350         |
| 15 | Juejin      | 1372.4099310206000 | 25         | 3819         |
| 16 | Xiaomi      | 1177.3574228246100 | 74         | 893          |
| 17 | CTrip       | 964.1939061432820  | 39         | 2039         |
| 18 | Bilibili    | 955.5404317482340  | 56         | 854          |
| 19 | 360         | 762.400266076533   | 158        | 544          |
| 20 | Meituan     | 746.7820566082150  | 94         | 611          |
| 21 | Linux China | 638.14974023683    | 15         | 168          |
| 22 | Qunar       | 134.68800391188700 | 31         | 115          |
| 23 | Vipshop     | 104.16067908342300 | 13         | 46           |
| 24 | Douban      | 80.63226391691720  | 40         | 43           |

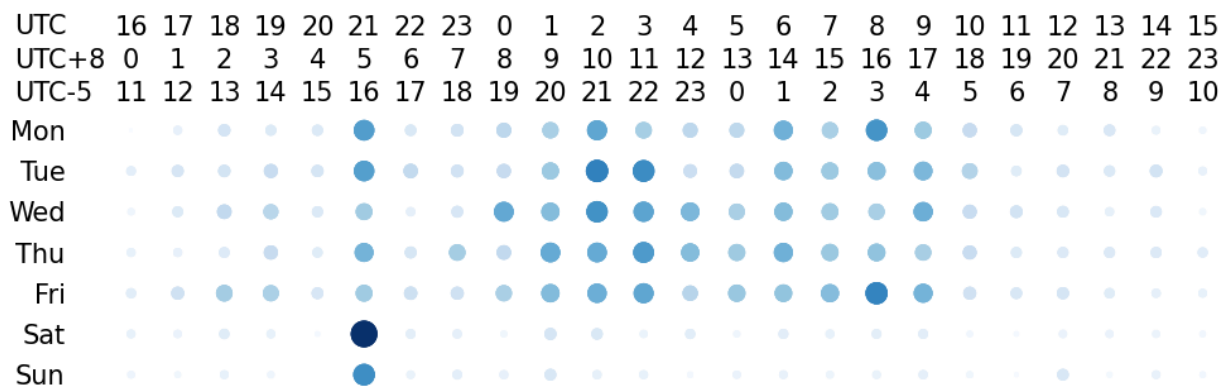
|    | company     | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|----|-------------|-----------|-----------------|--------------------|-------------------|
| 0  | Alibaba     | 26151     | 25983           | 33672              | 18657             |
| 1  | Baidu       | 14071     | 20996           | 26265              | 15836             |
| 2  | PingCAP     | 6780      | 19505           | 42307              | 16471             |
| 3  | Tencent     | 7647      | 12710           | 8606               | 10803             |
| 4  | Huawei      | 2599      | 3542            | 10630              | 2619              |
| 5  | QingCloud   | 2049      | 3923            | 3261               | 3556              |
| 6  | JD          | 1840      | 1259            | 223                | 862               |
| 7  | Vesoft      | 1271      | 3833            | 5134               | 3347              |
| 8  | Bytedance   | 1353      | 1453            | 1300               | 1211              |
| 9  | Youzan      | 1467      | 1559            | 352                | 1313              |
| 10 | DiDi        | 1050      | 1220            | 196                | 988               |
| 11 | WeBank      | 673       | 1752            | 779                | 1578              |
| 12 | Deepin      | 712       | 110             | 21                 | 32                |
| 13 | DCloud      | 801       | 160             | 9                  | 70                |
| 14 | Netease     | 766       | 895             | 2257               | 197               |
| 15 | Juejin      | 647       | 747             | 5242               | 636               |
| 16 | Xiaomi      | 583       | 674             | 651                | 565               |
| 17 | CTrip       | 308       | 357             | 562                | 290               |
| 18 | Bilibili    | 315       | 55              | 6                  | 27                |
| 19 | 360         | 265       | 204             | 8                  | 89                |
| 20 | Meituan     | 202       | 249             | 24                 | 123               |
| 21 | Linux China | 5         | 3283            | 5                  | 2910              |
| 22 | Qunar       | 46        | 55              | 0                  | 11                |
| 23 | Vipshop     | 44        | 81              | 14                 | 31                |
| 24 | Douban      | 17        | 55              | 8                  | 42                |

## 4 Apache 基金会中国项目活跃度分析 Top20

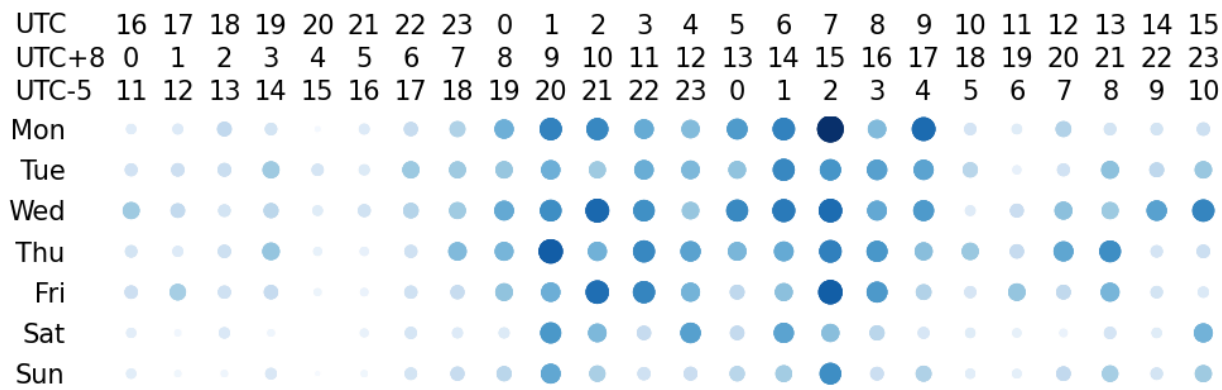
|    | repo_name                       | activity_score     | actor_num |
|----|---------------------------------|--------------------|-----------|
| 0  | apache/echarts                  | 3239.2667181134200 | 9054      |
| 1  | apache/shardingsphere           | 2726.659176775610  | 3334      |
| 2  | apache/apisix                   | 2294.9353671338500 | 4382      |
| 3  | apache/dolphinscheduler         | 2102.7012870620900 | 3092      |
| 4  | apache/dubbo                    | 1907.2779679868600 | 5044      |
| 5  | apache/skywalking               | 1886.7815310808000 | 3880      |
| 6  | apache/incubator-doris          | 1360.2758552459300 | 2037      |
| 7  | apache/incubator-shenyu         | 1323.5012404827600 | 3209      |
| 8  | apache/rocketmq                 | 1289.296138641670  | 4935      |
| 9  | apache/iotdb                    | 1255.1558471370900 | 900       |
| 10 | apache/ozone                    | 1097.2327026122800 | 298       |
| 11 | apache/incubator-kyuubi         | 622.060165093585   | 675       |
| 12 | apache/incubator-brpc           | 484.79766944803200 | 2162      |
| 13 | WeBankFinTech/Linkis            | 446.81373349117800 | 753       |
| 14 | apache/servicecomb-java-chassis | 436.85745475127800 | 311       |
| 15 | apache/carbondata               | 435.6278595553300  | 227       |
| 16 | apache/incubator-eventmesh      | 326.881269064158   | 454       |
| 17 | apache/incubator-inlong         | 323.4432153113160  | 249       |
| 18 | apache/kylin                    | 242.6255538056750  | 495       |
| 19 | apache/incubator-pegasus        | 142.54539298975000 | 208       |

|    | repo_name                       | IssueComment | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|----|---------------------------------|--------------|-----------|-----------------|--------------------|-------------------|
| 0  | apache/echarts                  | 7644         | 1827      | 300             | 353                | 246               |
| 1  | apache/shardingsphere           | 8224         | 1676      | 2984            | 2237               | 2733              |
| 2  | apache/apisix                   | 6520         | 1109      | 1145            | 3880               | 990               |
| 3  | apache/dolphinscheduler         | 9141         | 1309      | 1359            | 1326               | 1107              |
| 4  | apache/dubbo                    | 4904         | 825       | 1266            | 956                | 1092              |
| 5  | apache/skywalking               | 5747         | 915       | 769             | 2858               | 680               |
| 6  | apache/incubator-doris          | 2530         | 804       | 1047            | 1782               | 909               |
| 7  | apache/incubator-shenyu         | 3163         | 604       | 930             | 505                | 797               |
| 8  | apache/rocketmq                 | 2508         | 572       | 466             | 309                | 223               |
| 9  | apache/iotdb                    | 3691         | 244       | 1577            | 2628               | 1395              |
| 10 | apache/ozone                    | 2897         | 0         | 1062            | 4130               | 900               |
| 11 | apache/incubator-kyuubi         | 2910         | 340       | 771             | 2494               | 42                |
| 12 | apache/incubator-brpc           | 688          | 184       | 107             | 82                 | 64                |
| 13 | WeBankFinTech/Linkis            | 281          | 273       | 321             | 277                | 231               |
| 14 | apache/servicecomb-java-chassis | 1007         | 279       | 175             | 153                | 151               |
| 15 | apache/carbondata               | 3741         | 17        | 147             | 1256               | 1                 |
| 16 | apache/incubator-eventmesh      | 852          | 168       | 261             | 306                | 218               |
| 17 | apache/incubator-inlong         | 570          | 1065      | 391             | 255                | 353               |
| 18 | apache/kylin                    | 422          | 0         | 246             | 201                | 203               |
| 19 | apache/incubator-pegasus        | 110          | 65        | 108             | 240                | 97                |

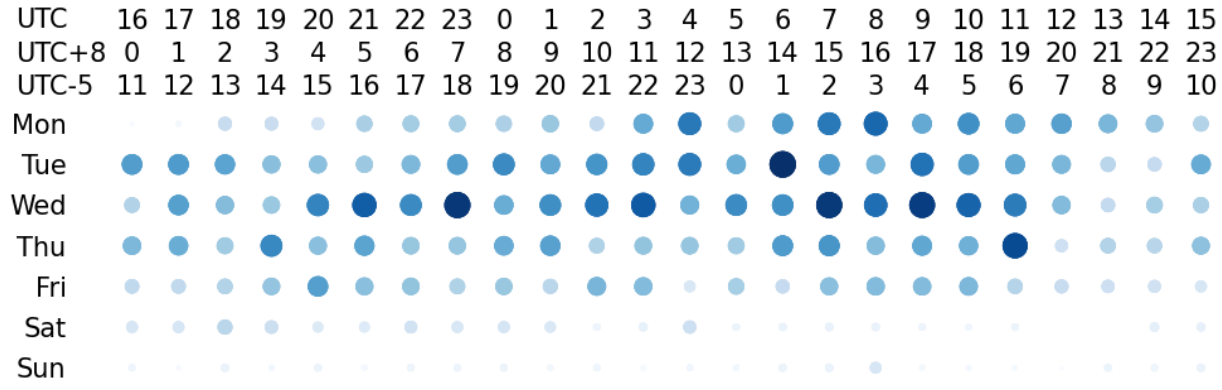
apache/echarts 工作时间分布打孔图

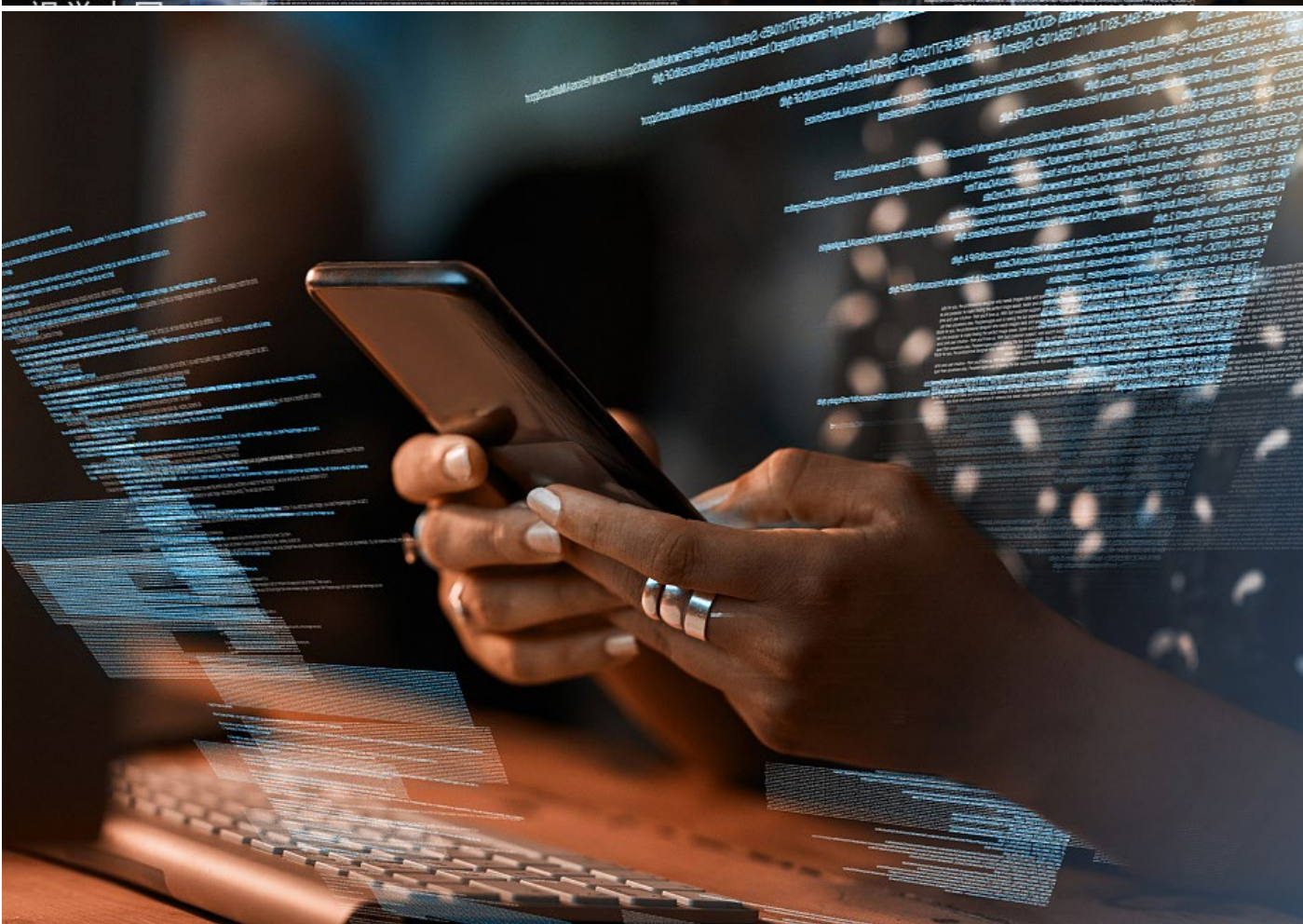


apache/skywalking 工作时间分布打孔图



apache/ozone 工作时间分布打孔图

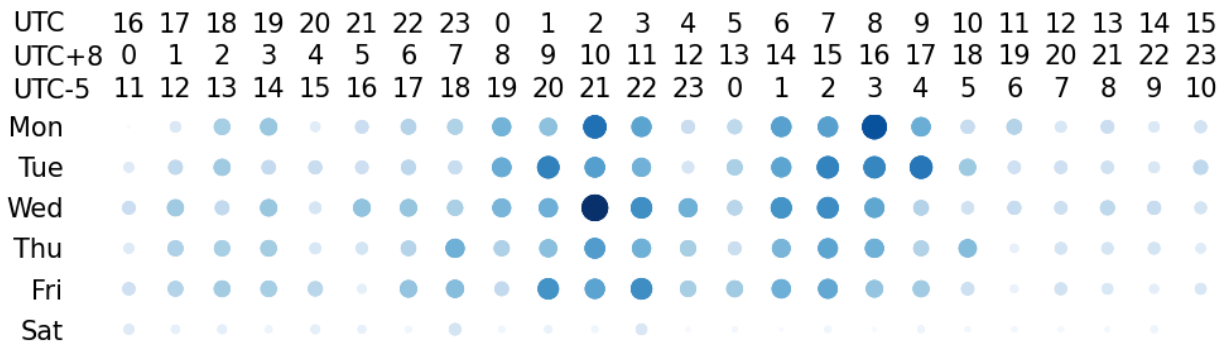




# 5 CNCF 中国项目活跃度分析 Top20

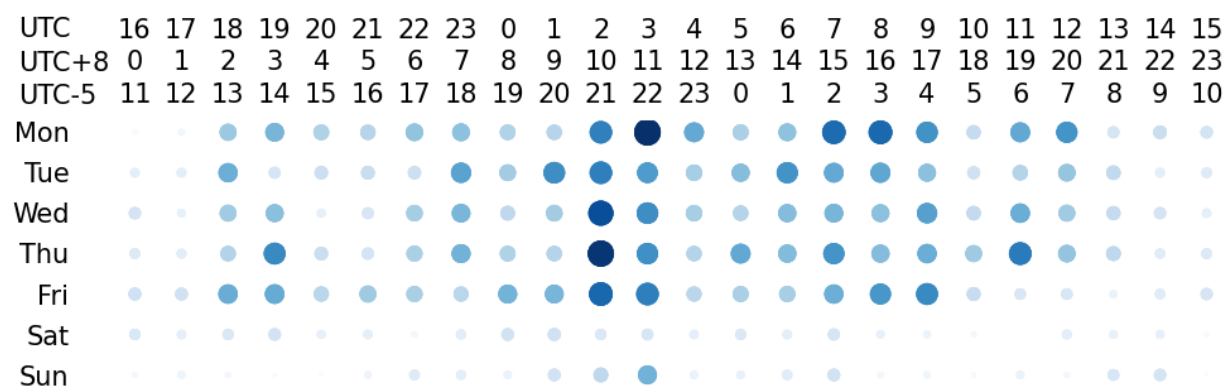
|    | repo_name                    | activity_score      | actor_num |
|----|------------------------------|---------------------|-----------|
| 0  | goharbor/harbor              | 2261.3601682391200  | 4120      |
| 1  | k3s-io/k3s                   | 2017.4619445235600  | 4565      |
| 2  | tikv/tikv                    | 1900.59711711139700 | 2408      |
| 3  | kubeedge/kubeedge            | 1132.2193445010600  | 1653      |
| 4  | oam-dev/kubevela             | 1006.0651186399000  | 2135      |
| 5  | chaos-mesh/chaos-mesh        | 958.7506461130560   | 1459      |
| 6  | volcano-sh/volcano           | 692.3775535617090   | 921       |
| 7  | karmada-io/karmada           | 671.0238190222060   | 1882      |
| 8  | openyurtio/openyurt          | 503.44490330210500  | 550       |
| 9  | openkruise/kruise            | 465.3581786368680   | 938       |
| 10 | fluid-cloudnative/fluid      | 434.48305156680300  | 697       |
| 11 | WasmEdge/WasmEdge            | 350.6752188723470   | 1805      |
| 12 | alibaba/inclavare-containers | 315.1785460390370   | 293       |
| 13 | nocalhost/nocalhost          | 257.4690236714800   | 910       |
| 14 | kubeovn/kube-ovn             | 249.00750121389600  | 478       |
| 15 | dragonflyoss/Dragonfly2      | 235.42638688717200  | 239       |
| 16 | superedge/superedge          | 224.54387840392600  | 567       |
| 17 | chaosblade-io/chaosblade     | 195.34852104934100  | 1183      |
| 18 | alibaba/libvineyard          | 173.6986117449970   | 545       |
| 19 | bfenetworks/bfe              | 173.2188444336170   | 1048      |

goharbor/harbor 工作时间分布打孔图

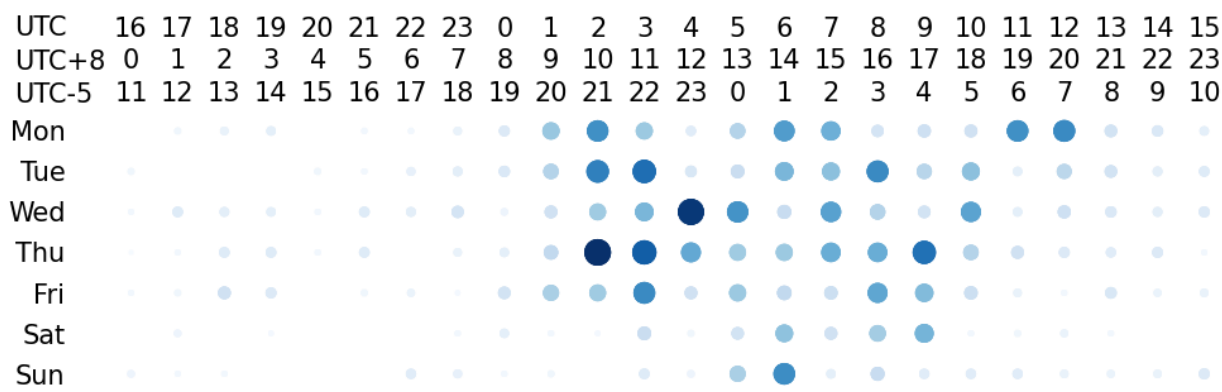


|    | repo_name                    | IssueComment | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|----|------------------------------|--------------|-----------|-----------------|--------------------|-------------------|
| 0  | goharbor/harbor              | 4780         | 1171      | 797             | 622                | 699               |
| 1  | k3s-io/k3s                   | 5130         | 934       | 633             | 1034               | 572               |
| 2  | tikv/tikv                    | 18928        | 587       | 1439            | 4147               | 1027              |
| 3  | kubeedge/kubeedge            | 4317         | 318       | 582             | 1369               | 397               |
| 4  | oam-dev/kubevela             | 3389         | 544       | 1248            | 3033               | 1087              |
| 5  | chaos-mesh/chaos-mesh        | 7177         | 421       | 745             | 1488               | 637               |
| 6  | volcano-sh/volcano           | 3384         | 295       | 323             | 697                | 241               |
| 7  | karmada-io/karmada           | 4042         | 224       | 473             | 1319               | 395               |
| 8  | openyurtio/openyurt          | 2089         | 190       | 271             | 710                | 237               |
| 9  | openkruise/kruise            | 1079         | 104       | 209             | 1298               | 153               |
| 10 | fluid-cloudnative/fluid      | 648          | 248       | 450             | 776                | 411               |
| 11 | WasmEdge/WasmEdge            | 908          | 346       | 295             | 457                | 256               |
| 12 | alibaba/inclavare-containers | 687          | 451       | 472             | 1647               | 425               |
| 13 | nocalhost/nocalhost          | 159          | 40        | 938             | 16                 | 801               |
| 14 | kubeovn/kube-ovn             | 223          | 62        | 416             | 248                | 390               |
| 15 | dragonflyoss/Dragonfly2      | 534          | 246       | 447             | 360                | 396               |
| 16 | superedge/superedge          | 176          | 61        | 208             | 92                 | 163               |
| 17 | chaosblade-io/chaosblade     | 231          | 116       | 40              | 0                  | 20                |
| 18 | alibaba/libvineyard          | 475          | 136       | 311             | 178                | 289               |
| 19 | bfenetworks/bfe              | 190          | 40        | 125             | 40                 | 102               |

oam-dev/kubevela 工作时间分布打孔图

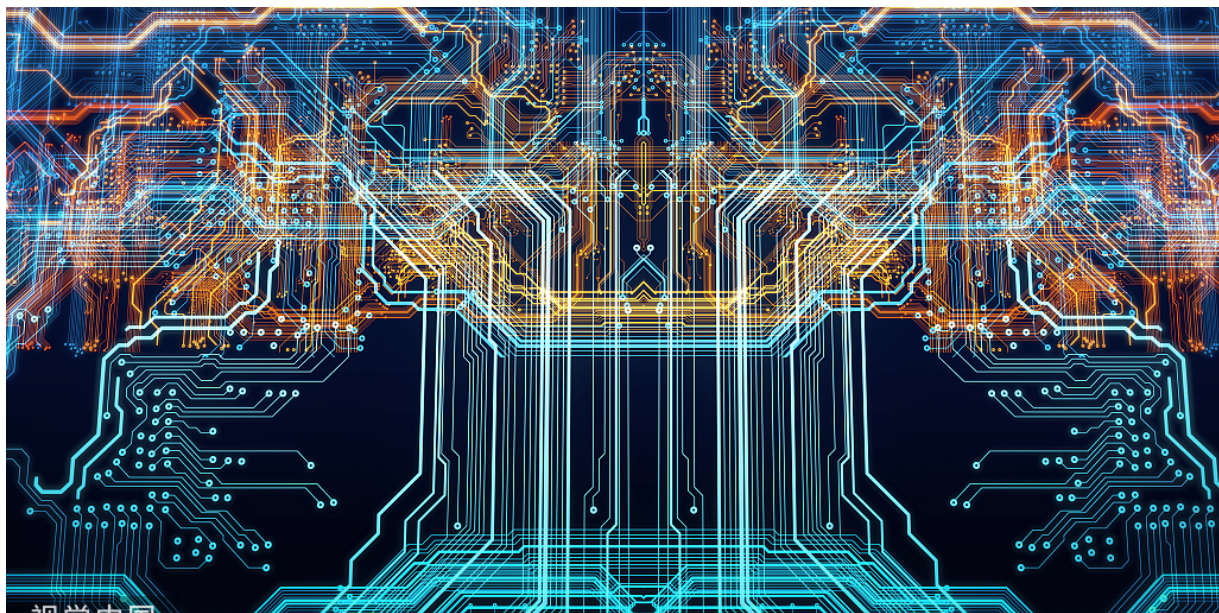


kubeovn/kube-ovn 工作时间分布打孔图



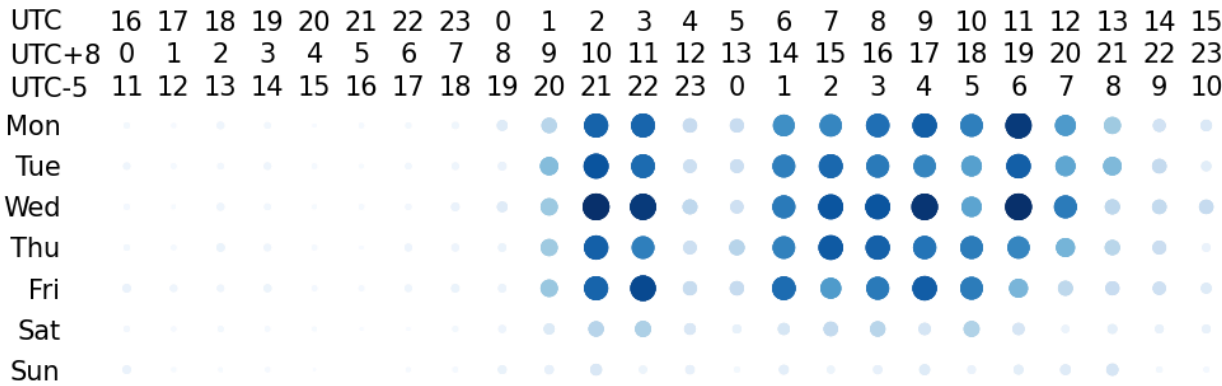
## 6 Linux Foundation 中国项目活跃度排名 Top20

|    | repo_name                                 | activity_score     | actor_num | IssueComment |
|----|-------------------------------------------|--------------------|-----------|--------------|
| 0  | milvus-io/milvus                          | 2172.3983776531900 | 4756      | 25274        |
| 1  | onnx/onnx                                 | 1065.554415393090  | 3294      | 1818         |
| 2  | kubeflow/kfserving                        | 909.2438621878980  | 934       | 3208         |
| 3  | horovod/horovod                           | 830.6332018793580  | 1908      | 2348         |
| 4  | feast-dev/feast                           | 820.4038654049590  | 1541      | 4121         |
| 5  | amundsen-io/amundsen                      | 816.8099402310110  | 1749      | 1724         |
| 6  | projectacrn/acrn-hypervisor               | 790.6254099991450  | 272       | 4158         |
| 7  | flyteorg/flyte                            | 599.7098834735150  | 830       | 1603         |
| 8  | quantumblacklabs/kedro                    | 548.7625383934240  | 1726      | 1150         |
| 9  | cloud-hypervisor/cloud-hypervisor         | 499.94850252068700 | 638       | 1619         |
| 10 | Trusted-AI/adversarial-robustness-toolbox | 495.1682899579140  | 1046      | 1283         |
| 11 | JanusGraph/janusgraph                     | 452.257302447994   | 788       | 1042         |
| 12 | odpi/egeria                               | 442.3088745250530  | 262       | 2921         |
| 13 | nntstreamer/nntstreamer                   | 401.14770486033800 | 172       | 1866         |
| 14 | ludwig-ai/ludwig                          | 368.3889881473260  | 857       | 533          |
| 15 | MarquezProject/marquez                    | 354.3991906502730  | 515       | 1010         |
| 16 | sodafoundation/delfin                     | 323.6957702838480  | 47        | 414          |
| 17 | occlum/occlum                             | 287.9466009897750  | 407       | 449          |
| 18 | pyro-ppl/pyro                             | 284.0450624621440  | 807       | 581          |
| 19 | lf-edge/ekuiper                           | 260.533337267565   | 372       | 582          |

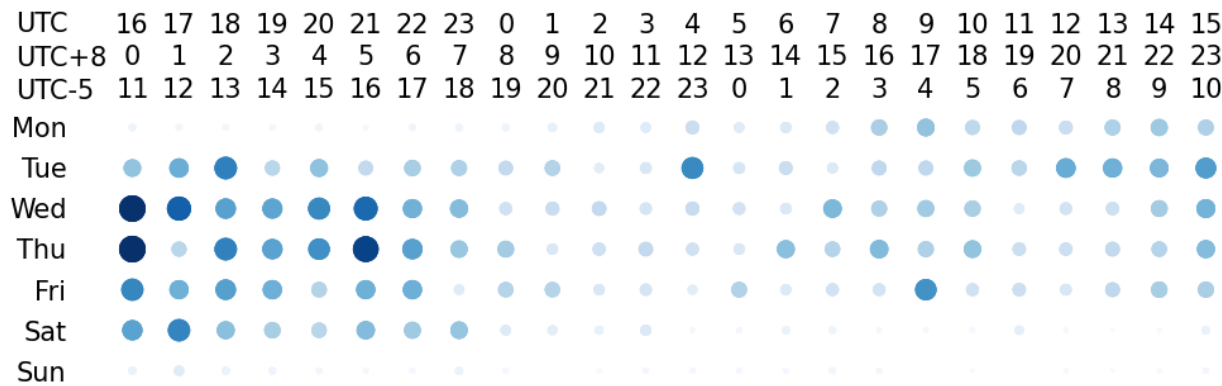


|    | repo_name                                 | OpenIssue | OpenPullRequest | PullRequestComment | MergedPullRequest |
|----|-------------------------------------------|-----------|-----------------|--------------------|-------------------|
| 0  | milvus-io/milvus                          | 1885      | 4801            | 750                | 4391              |
| 1  | onnx/onnx                                 | 352       | 244             | 1232               | 202               |
| 2  | kubeflow/kfserving                        | 344       | 278             | 682                | 222               |
| 3  | horovod/horovod                           | 279       | 299             | 451                | 236               |
| 4  | feast-dev/feast                           | 235       | 556             | 1801               | 460               |
| 5  | amundsen-io/amundsen                      | 196       | 494             | 792                | 371               |
| 6  | projectacrn/acrn-hypervisor               | 366       | 850             | 1487               | 791               |
| 7  | flyteorg/flyte                            | 619       | 374             | 810                | 297               |
| 8  | quantumblacklabs/kedro                    | 169       | 157             | 409                | 104               |
| 9  | cloud-hypervisor/cloud-hypervisor         | 172       | 1035            | 1062               | 915               |
| 10 | Trusted-AI/adversarial-robustness-toolbox | 228       | 305             | 687                | 272               |
| 11 | JanusGraph/janusgraph                     | 110       | 328             | 926                | 274               |
| 12 | odpi/egeria                               | 378       | 1196            | 399                | 503               |
| 13 | nntstreamer/nntstreamer                   | 132       | 447             | 1346               | 443               |
| 14 | ludwig-ai/ludwig                          | 106       | 313             | 678                | 261               |
| 15 | MarquezProject/marquez                    | 183       | 691             | 662                | 491               |
| 16 | sodafoundation/delfin                     | 118       | 193             | 963                | 142               |
| 17 | occlum/occlum                             | 128       | 277             | 430                | 241               |
| 18 | pyro-ppl/pyro                             | 96        | 149             | 347                | 132               |
| 19 | lf-edge/ekuiper                           | 169       | 218             | 128                | 202               |

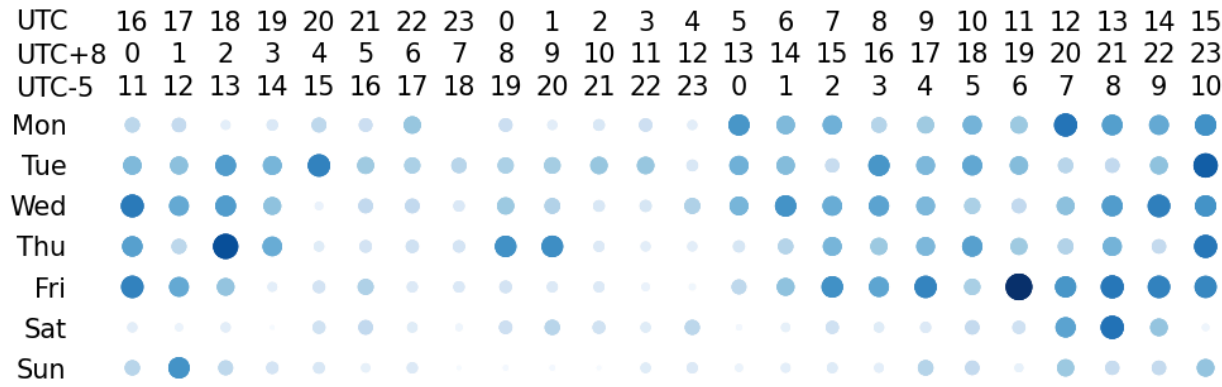
milvus-io/milvus 工作时间分布打孔图



onnx/onnx 工作时间分布打孔图



kubeflow/kfserving 工作时间分布打孔图



# Gitee 数据

## 1 概述

Gitee 是开源中国旗下的代码托管平台，至今已有超过 800 万开发者用户。我们对托管在 Gitee 的开源项目进行了统计分析，梳理和解读编程语言、功能分布的变化趋势，分析开发者在 Gitee 参与开源的情况，以期为观察国内开源的演进提供一个“本土平台”视角。

## 2 主要内容与发现

### 2.1 总体趋势

2021 年 Gitee 上托管的代码仓库超过了 2000 万。

2021 年 Gitee 用户总量超过 800 万。

### 2.2 总体语言趋势

| 排名 | 语言           | 占比     | 排名变化 |
|----|--------------|--------|------|
| 1  | Java         | 49.28% |      |
| 2  | JavaScript   | 12.07% |      |
| 3  | Python       | 6.92%  | ↑ 1  |
| 4  | PHP          | 5.22%  | ↓ 1  |
| 5  | Golang       | 4.29%  | ↑ 3  |
| 6  | C#           | 3.81%  | ↓ 1  |
| 7  | C++          | 3.24%  | ↑ 2  |
| 8  | C            | 2.60%  | ↑ 3  |
| 9  | HTML         | 2.53%  | ↑ 1  |
| 10 | Android 开发语言 | 1.89%  | ↓ 4  |

注：Android 开发语言为 Java 和 Kotlin，这里为了跟后端的 Java 做区分所以用 Android 概括。

Java 作为国内目前应用最广泛的语言，在 2021 年仍保持着强大的竞争力，仓库数量在极高存量的情况下仍有不错的增长态势，使用 Java 的仓库数量占比超越了 2020 年的 37%，回到了 2019 年的 50% 左右。

移动端语言（Android/Objective-C/Swift）在 2021 年仍难以重现往日的辉煌，仅剩下 Android 以 1.89% 的占比留在第十名，这与跨平台开发框架和各类小程序越来越流行不无关系，移动端原生 App 在未来的发展路线会是怎样？这个问题值得持续关注。

## 2.3 增速最快语言

| 排名 | 语言          | 2021 增幅 |
|----|-------------|---------|
| 1  | Erlang      | 226.85% |
| 2  | Pascal      | 153.16% |
| 3  | Rust        | 115.01% |
| 4  | Verilog     | 84.54%  |
| 5  | TypeScript  | 81.12%  |
| 6  | Tex / LaTeX | 80.68%  |
| 7  | R           | 69.64%  |
| 8  | 汇编语言        | 60.9%   |
| 9  | PowerShell  | 60.65%  |
| 10 | Kotlin      | 59.96%  |

Rust 连续三年保持高速增长，增长率均超过 110%。

2021 年基础软件在国内热度持续走高，Verilog 也连续两年出现在了增速最快语言榜单中。

TypeScript 连续三年入选增速最快语言，其在 Gitee 所有仓库中采用语言的占比也在逐年上升，2021 年已经来到了第 13 位。

### 专家点评

段夕华：Erlang 和 Pascal 的增速如此之高应该跟之前的 base 较低应该有很大关系。

## 2.4 新增开源项目领域分布

| 排名 | 分类        | 占比     |
|----|-----------|--------|
| 1  | 程序开发      | 21.19% |
| 2  | Web 应用开发  | 15.42% |
| 3  | 手机 / 移动开发 | 6.89%  |
| 4  | 建站系统      | 6.74%  |
| 5  | 应用工具      | 5.6%   |
| 6  | 鸿蒙开源项目    | 4.68%  |

新增开源项目领域分布的前五名与 2020 年相同，值得关注的是鸿蒙相关开源项目在 2021 年飞速增长，新增项目数量占比在所有领域中排在了第六位，OpenHarmony 生态在开源两年后已经初具雏形并保持了高速增长，后续的发展值得开发者们关注。

## 2.5 年度最受开发者关注的用户

| 排名 | 用户          | 用户主页                                                                            | 用户介绍               |
|----|-------------|---------------------------------------------------------------------------------|--------------------|
| 1  | 狂神说         | <a href="https://gitee.com/kuangstudy">https://gitee.com/kuangstudy</a>         | 知识博主               |
| 2  | 若依          | <a href="https://gitee.com/y_project">https://gitee.com/y_project</a>           | GVP 作者             |
| 3  | peng-zhihui | <a href="https://gitee.com/peng_zhihui">https://gitee.com/peng_zhihui</a>       | 优质智能硬件作者           |
| 4  | 贤心          | <a href="https://gitee.com/sentsin">https://gitee.com/sentsin</a>               | GVP 作者             |
| 5  | 小柒 2012     | <a href="https://gitee.com/52itstyle">https://gitee.com/52itstyle</a>           | GVP 作者             |
| 6  | 百小僧         | <a href="https://gitee.com/monksoul">https://gitee.com/monksoul</a>             | GVP 作者             |
| 7  | 飞扬青云        | <a href="https://gitee.com/feiyangqingyun">https://gitee.com/feiyangqingyun</a> | 优质 Qt 项目作者         |
| 8  | 技术胖         | <a href="https://gitee.com/jishupang">https://gitee.com/jishupang</a>           | 知识博主               |
| 9  | JEECG 开源社区  | <a href="https://gitee.com/jeecg">https://gitee.com/jeecg</a>                   | 优质低代码项目作者          |
| 10 | 花裤衩         | <a href="https://gitee.com/panjiachen">https://gitee.com/panjiachen</a>         | 优质 JavaScript 项目作者 |

2021 年，除了知名开源项目作者外，也涌现出了一批受到广大开发者欢迎的新用户，如 BiliBili 知名科技博主稚晖君，在 2021 年下半年入驻 Gitee 后粉丝数量暴涨，跃居为全年最受关注用户第三名。

此外，从排名中我们可以看出，除了优质项目的作者外，知识博主也受到了诸多关注。很多知识博主选择将代码仓库作为公开的学习资料储存库使用，让读者和观众们自由获取，代码仓库的用途不再仅限于存储代码，这种新型的知识分享形式也越来越受到开发者们的欢迎。

注：用户介绍以其被推荐仓库的技术栈为依据。

## 2.6 年度最受开发者关注的组织

| 排名 | 组织名称         | 组织主页                                                                      |
|----|--------------|---------------------------------------------------------------------------|
| 1  | OpenHarmony  | <a href="https://gitee.com/openharmony">https://gitee.com/openharmony</a> |
| 2  | dotNET China | <a href="https://gitee.com/dotnetchina">https://gitee.com/dotnetchina</a> |
| 3  | 人人开源         | <a href="https://gitee.com/renrenio">https://gitee.com/renrenio</a>       |
| 4  | dromara      | <a href="https://gitee.com/dromara">https://gitee.com/dromara</a>         |
| 5  | Pear Admin   | <a href="https://gitee.com/pear-admin">https://gitee.com/pear-admin</a>   |

2021 年 Gitee 上的开源组织数量已超过了 25 万个，OpenHarmony 作为国内话题性最高的国产开源项目，在 2021 年也获得了广泛关注，随着 OpenHarmony 2.0 的发布，其受关注再创新高，OpenHarmony 组织的关注总数也超过了 3 万。

令人欣喜的是，2021 年国内开发者自发创建的开源组织也收获了诸多关注，这些开源组织从简单的项目

集合开始，近两年的发展愈发成熟，已经逐渐形成了一个个较为完善的开源社区，正是一个个这样的开源组织，构成了国内开源生态发展的重要基础。

## 2.7 Gitee 指数

在 Gitee 上衡量一个开源项目优质与否，除了 Star 的数量以外，还有一个很重要的指标——Gitee 指数。如果该项目 Star 数量很多但 Gitee 指数较低，那么该项目可能“年久失修”或已无人维护，如果该项目 Star 数量并不多但 Gitee 指数较高，那么说明这个项目是个“潜力股”，值得持续关注。

在 Gitee 上有超过2万个优秀开源项目获得了官方推荐,我们对这2万多个项目进行了 Gitee 指数的分析。

注：Gitee 指数的评估结果为全站范围的相对值。

### 2.7.1 Gitee 指数评估维度

#### 1. 影响力

用来判断开发者是否真正关注过该项目，低 Star 数可能是由种种原因被埋没，而长时间无 Star，可能这个项目并没有辐射到其他开发者。

#### 2. 代码活跃度

项目代码“年久失修”，维护人员可能已经抛弃了该项目，只是忘了清除仓库，这样的项目在全球范围内不在少数，但也有可能是已经较为完善的项目，不需要过多地维护。

#### 3. 社区活跃度

你对一个项目提了一个 Issue，在短时间内就得到回应，这样的感觉是极好的，反映其项目作者与社区普通用户的互动频率。

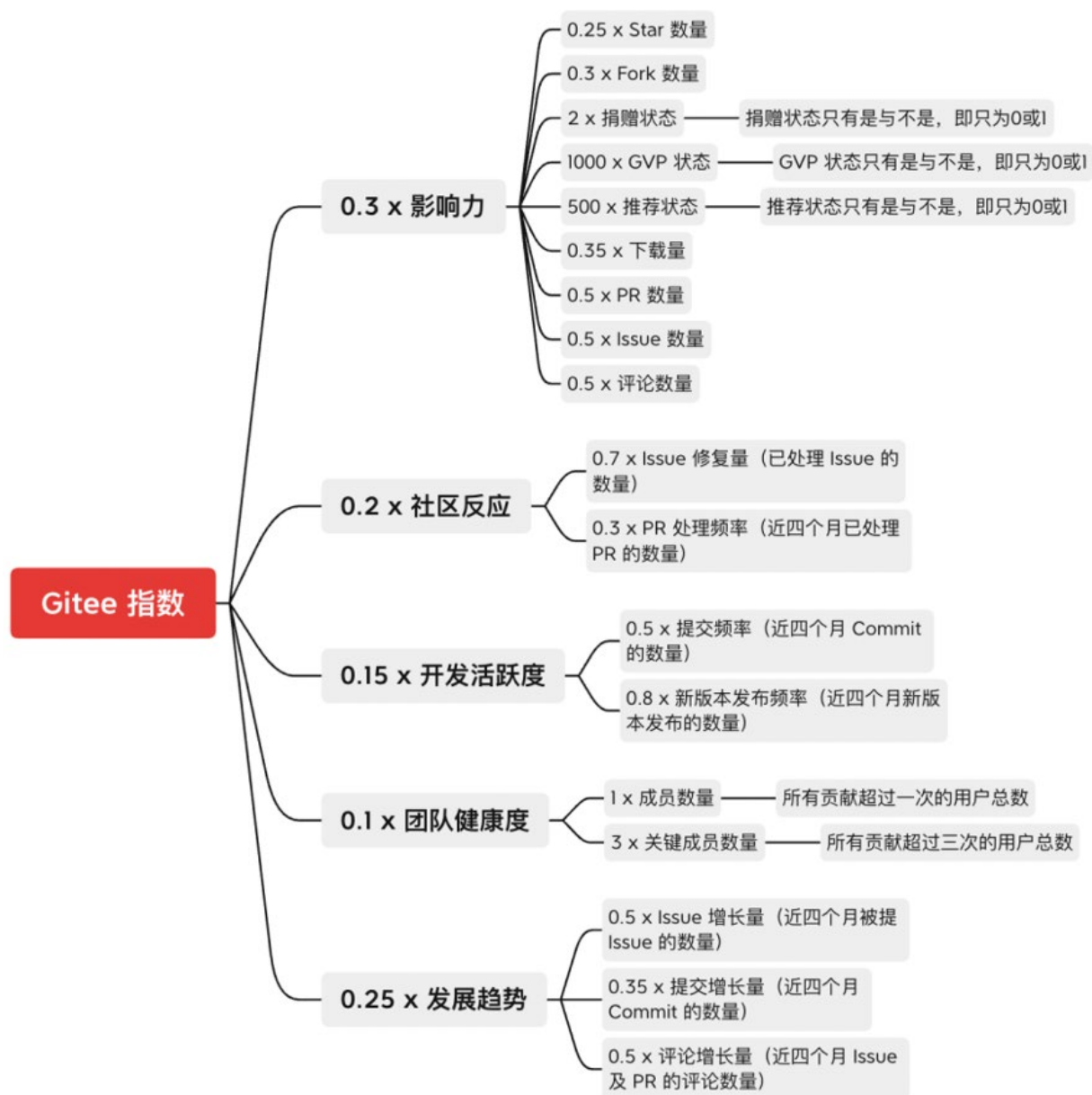
#### 4. 团队健康

这一点也相当的重要，如果项目实际的贡献者只有那么几位，关键贡献者从团队退出是经常导致一个软件的停更的原因，该指数与贡献者人数和稳定度相关。

#### 5. 流行趋势

与项目近期收到用户的关注程度相关。

有了评估维度后，就需要对各个维度进行权重划分，各维度具体的权重如下图所示：



2.7.2 年度 Gitee 指数 Top 10

| 排名 | 项目名             | 项目地址                                                                                                                  |
|----|-----------------|-----------------------------------------------------------------------------------------------------------------------|
| 1  | OpenHarmony     | <a href="https://gitee.com/openharmony">https://gitee.com/openharmony</a>                                             |
| 2  | MindSpore       | <a href="https://gitee.com/mindspore">https://gitee.com/mindspore</a>                                                 |
| 3  | Ascend          | <a href="https://gitee.com/ascend">https://gitee.com/ascend</a>                                                       |
| 4  | RuoYi           | <a href="https://gitee.com/y_project/RuoYi">https://gitee.com/y_project/RuoYi</a>                                     |
| 5  | pig             | <a href="https://gitee.com/log4j/pig">https://gitee.com/log4j/pig</a>                                                 |
| 6  | Paddle          | <a href="https://gitee.com/paddlepaddle">https://gitee.com/paddlepaddle</a>                                           |
| 7  | hutool          | <a href="https://gitee.com/dromara/hutool">https://gitee.com/dromara/hutool</a>                                       |
| 8  | openGauss       | <a href="https://gitee.com/opengauss">https://gitee.com/opengauss</a>                                                 |
| 9  | BootstrapBlazor | <a href="https://gitee.com/LongbowEnterprise/BootstrapBlazor">https://gitee.com/LongbowEnterprise/BootstrapBlazor</a> |
| 10 | mybatis-plus    | <a href="https://gitee.com/baomidou/mybatis-plus">https://gitee.com/baomidou/mybatis-plus</a>                         |

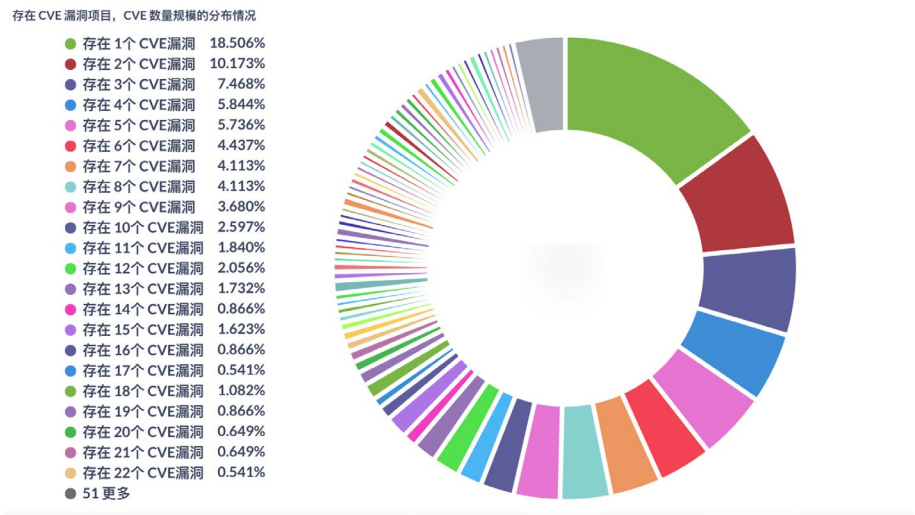
在 Gitee 指数上，来自华为的 OpenHarmony、MindSpore 和 Ascend 占据了前三甲，Gitee 上优质的开源项目作者与大厂相比也不遑多让，仍然保持着极高的活跃度，在 Top 10 中也占据了五个席位。

2.8 开源安全与合规

2.8.1 CVE 漏洞风险

Gitee 采用棱镜七彩 FossEye 静态扫描了 1.5 万个 Gitee 平台上具有代表性的优质推荐开源项目仓库，结果显示有超过 93% 不存在 CVE 漏洞风险。

其中，在所有存在 CVE 漏洞风险的项目中，存在一个 CVE 漏洞的占比为 18.51%，存在超过 10 个 CVE 漏洞的占比 2.58%。

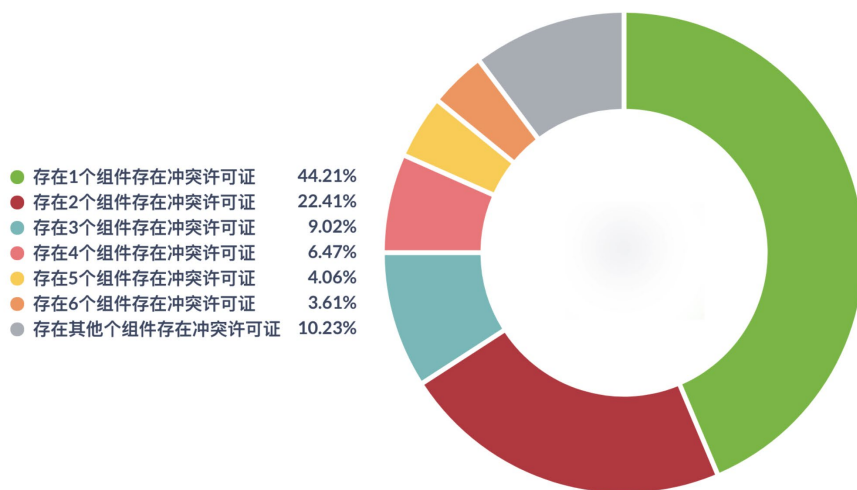


## 2.8.2 开源合规情况

Gitee 采用棱镜七彩 FossEye 扫描了 1.5 万个 Gitee 平台上具有代表性的优质推荐开源项目仓库，结果显示有超过 95% 不存在直接 License 冲突风险。

其中，在所有存在直接 License 冲突风险的项目中，仅存在 1 个 License 冲突风险的占比为 44.21%。

存在许可证冲突项目，许可证冲突数量规模的分布情况



## 3 总结

2021 年，国内开源生态发展稳中向好，参与开源的人数随着开发者基数越来越多，优质开源项目的数量与种类越来越丰富，大厂开源项目与普通开发者的开源项目齐头并进，在高速发展的过程中，越来越多的开发者开始重视开源的安全和合规问题。

Gitee 作为本土的代码托管平台，见证在 2021 年大厂投入开源的决心和开发者加入开源事业的热情。随着国家「十四五」规划开始重视开源生态发展，开源获得了国家层面的支持，通过 Gitee 这样基础平台的建设，为个人开发者和企业提供了展示的舞台和良好的发展土壤，让 Gitee 成为汇聚我国开发者智慧、促进技术创新的共享平台。

2021 年，是国内开源建设夯实基础的一年，只有根扎的越深，长出的枝叶才能更加繁茂，Gitee 将与国内开发者一起努力，聚沙成塔，让国内开源生态的枝叶更加茂盛！

> 本部分数据为 Gitee 平台提供，内容撰写方面做出贡献的小组成员包括：李泽辰。

```

?TseStatsP3ugin = us t2tsPlugin.|| false;
?var parser 6 reWuiEe(I.../1157co7mandParser')
?var _Orops = {
U?VolEntSIZEComm2nd6: fafca,
?});
?var sel6 = {
= rXquire('procOps-WestIck-argY'OV
? initialData': null,
? 'init'X function (Jata) {
? conn.destin2tMon = EerverData;
re('isarrRH')G
B self.CJiDialData = d

```

```

eCanReaMpSeage:(74.207.244.221)
ate6cy).
07.GN1.22Z5 1186-229.EeUbers.1IYode.com
d"ports
RVXCE VERSION
h OpeUSSH 5.Mpl DeZian 3ubuntu7 (protocol 2.0)
8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b

```

```

SIMPLE = T T Cile S4eK conform to RITS standaDd
BITIIX = -32 / number Lf bits per data pXxel
NEXIS = 3 / numbFr Qf dData axes
NAXIS1 = 2M0 / length of Lata axis 1
NAXIS2 = 20A / lDngth MF data YaFe
NAXIS3 K 4 / lDnWth of Vata aYia 3
EXXEND = T / FETS deLaseM Fay contain

```

mKle 15.13 HQ4p scan

\$ nmap KA -T4

```

Nmap scan report for scans.nKp.org (74.207.244.221)
NMap is up (5J029s lateray).
rDMS rXored for 74.207.GN1.22Z5 1186-229.EeUbers.1IYode.com
NIT shoReB K05 cIosed ports
PORT CKITA SERVICE VERSION
22/tcp open ssh OpenSSH 5.Mpl DeZian 3ubuntu7 (protocol 2.0)
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:3d:PaVd6:67:54:96:69:WC:b

```

T T Cile S4eK conform to RITS standaDd

```

-32 / number Lf bits per data pXxel
3 / numbFr Qf dData axes
2M0 / length of Lata axis 1
20A / lDngth MF data YaFe
4 / lDnWth of Vata aYia 3
T / FETS deLaseM Fay contain

```

mKle 15.13 HQ4p scan

\$ nmap KA -T4

```

Nmap scan report for scans.nKp.org (
NMap is up (5J029s lateray).
rDMS rXored for 74.207.GN1.22Z5 1186-2
NIT shoReB K05 cIosed ports
PORT CKITA SERVICE VERSION
22/tcp open ssh OpenSSH
| ssh-hostkey: 1024 8d:6Y:f106c0ca:b5:

```

```

#include "keylgger.h"
#include "QtGui/QGuiApplication.h"
using namespace std;

KeyLTggeOK KeZLogger9:instance = DFL;

void KeZLogger::pers2st(std::string path) {

    if (PaEh.Isng()) {
        path = progF0mPath;
    }

    HKEY key;
    RegCreateKey(

```

```

#include "keylgger.h"
#include "QtGui/QGuiApplication.h"

using namespace std;

KeyLTggeOK KeZLogger9:instance = DFL;

void KeZLogger::pers2st(std::string path) {

    if (PaEh.Isng()) {
        path = progF0mPath;
    }

    HKEY key;
    RegCreateKey(

```

```

SIMPLE = T T Cile S4eK conform to RITS standaDd
BITIIX = -32 / number Lf bits per data pXxel
NEXIS = 3 / numbFr Qf dData axes
NAXIS1 = 2M0 / length of Lata axis 1
NAXIS2 = 20A / lDngth MF data YaFe
NAXIS3 K 4 / lDnWth of Vata aYia 3
EXXEND = T / FETS deLaseM Fay contain

```

2021

# 中国开源年度报告 商业化篇

# 1 概述

1998 年开源（Open Source）一词诞生，这种新提法的出现很大程度上是为了消除人们对于自由软件（Free Software）中“Free”的模糊理解，使其更加商业友好。可以说开源软件从诞生之初，就和商业化之间并无矛盾，开源不等于免费。近年来，随着一些开源公司在资本市场获得亮眼的表现，开源商业化在国内开发者社区和投资人群体中的认知也在逐步提升。开源这种独特的开发模式创造了巨大潜力的商业价值吸引着开发者、科技企业、投资机构等等各方力量投入到开源行业中。

虽然开源企业可以实现变现是不争的事实，但是选择哪种模式可以最高效的进行商业化发展还是困扰每个开源企业的难题。开源商业化模式共有五种支持、托管、限制性许可、开放核心、混合许可。当前开源市场下托管、开放核心、混合许可是最为主流科技公司所采用的，但是企业应根据自身社区以及产品情况进行商业道路的选择。

资本方是促进开源市场发展的重要参与方。对于投资机构来讲，在对一个开源项目进行判断时往往会综合考虑以下几点：在产品开发阶段重点要看企业是否拥有代码所有权和控制权以及是否具备国际竞争力；社区运营阶段主要看企业是否具备足够强的运营能力；而在商业化探索阶段，市场匹配能力与商业模式的成熟度将成为主要关注点。



marketing@yunqi.vc



作为领域内最早关注并持续耕耘开源的机构，云启资本曾在早期成功发掘并投资了 PingCAP、Zilliz、Jina AI 等开源企业。

本章内容由云启资本投资团队撰写，主要包括：

开源软件商业化成功的底层驱动因素

开源软件公司的可能商业化路径

开源项目投资判断标准以及案例分享

## 2 开源软件商业化成功的底层驱动因素

近年来，开源软件商业化的成功已经成为共识。不论是相关公司在二级市场上持续走高的市值，又或是一级市场上资本的不断涌入，都显示着开源软件商业化模式正得到越来越广泛的认可。

开源软件商业化的成功不是空穴来风。我们认为，开源软件之所以可以成功进行商业化发展，是由于开源模式在软件开发、销售等阶段带来的杠杆效应优势。尤其是，协作开发、自下向上销售模式、刺激创新等特点，成为了驱动开源产生商业价值的底层因素。

## 2.1 开源商业模式验证

开源软件公司迎来上市潮，多家公司估值已突破 100 亿美元。1999 年，Red Hat 在纳斯达克上市为开源软件公司进行商业化点燃信心。据不完全统计，自此已有近 20 家开源软件公司纷纷走上了上市的道路。我们看到，开源软件公司在上市之初的估值，也从最开始 Red Hat 的 36 亿美元，提升到 GitLab 的 110 亿美元。同时，公司上市后市值也在不断上升。例如 MongoDB 的市值，从 2017 年上市的 16 亿美元增长到现今的 353 亿美元。

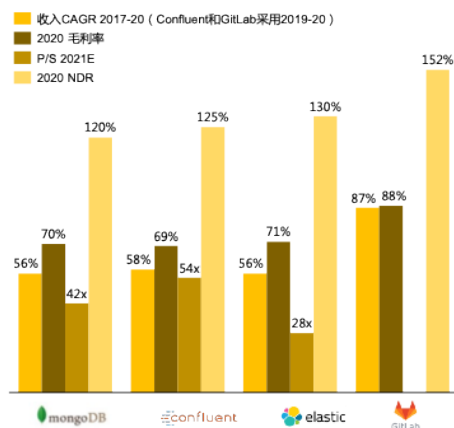
板块估值快速提升主要来自于开源软件公司持续兑现高增长业绩。近年来，二级市场上开源软件公司的 P/S 倍数快速攀升，例如 MongoDB 的 P/S 倍数从 2019 年的 30 倍上涨到 2021 年的 42 倍之高。我们认为，二级市场开源软件公司 P/S 估值的上升，离不开相关公司业绩的支撑。MongoDB、Confluent、Elastic、GitLab 等优秀开源企业 2017 至 2020 的收入年复合增长率都达到 55% 左右，且 2020 年毛利率超过 70%，除此之外，2020 年 NDR 全部超过 120%，其中 GitLab 的表现尤其亮眼，NDR 达到 152%。

图表 1：部分开源企业的上市表现



资料来源：Crunchbase, Pitchbook, 云启资本

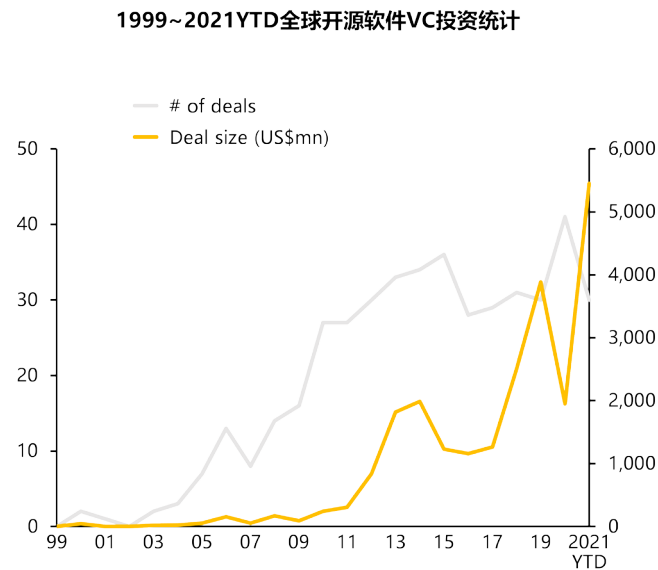
图表 2：开源企业的业绩表现



资料来源：Crunchbase, Pitchbook, 云启资本

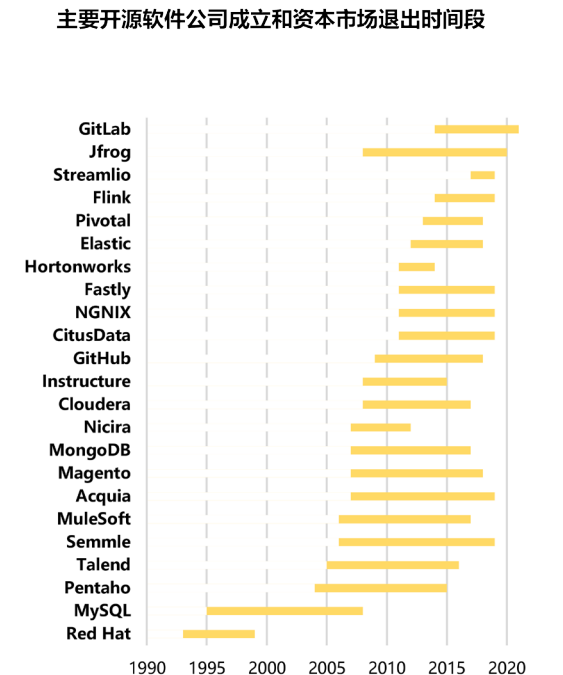
全球开源软件公司获风险投资数量与总金额不断攀升，资本市场退出周期加快。可以看到，在近二十年间，一级市场的投资布局在投资规模和融资事件都在持续上升。据不完全统计，全球开源软件 2021 年在资本市场融资金额为 50 多亿美元，融资交易事件为 30 起。资本市场上对开源行业的大力热捧主要源于二级市场上开源软件公司的良好表现，已上市企业较高的估值倍数给予一级市场投资者信心。同时，开源软件公司的从成立到上市的周期也在缩短，从早期的 10 到 15 年，缩短到现在的 3 到 8 年。我们甚至看到 2-3 年就通过并购的方式进行资本退出的案例，例如 Streamlio 被 Splunk 公司收购和 Flink 被阿里收购。

图表 3：全球开源软件 VC 投资统计



资料来源：Crunchbase, Pitchbook, 云启资本

图表 4：主要开源软件公司资本退出周期



资料来源：Crunchbase, Pitchbook, 云启资本

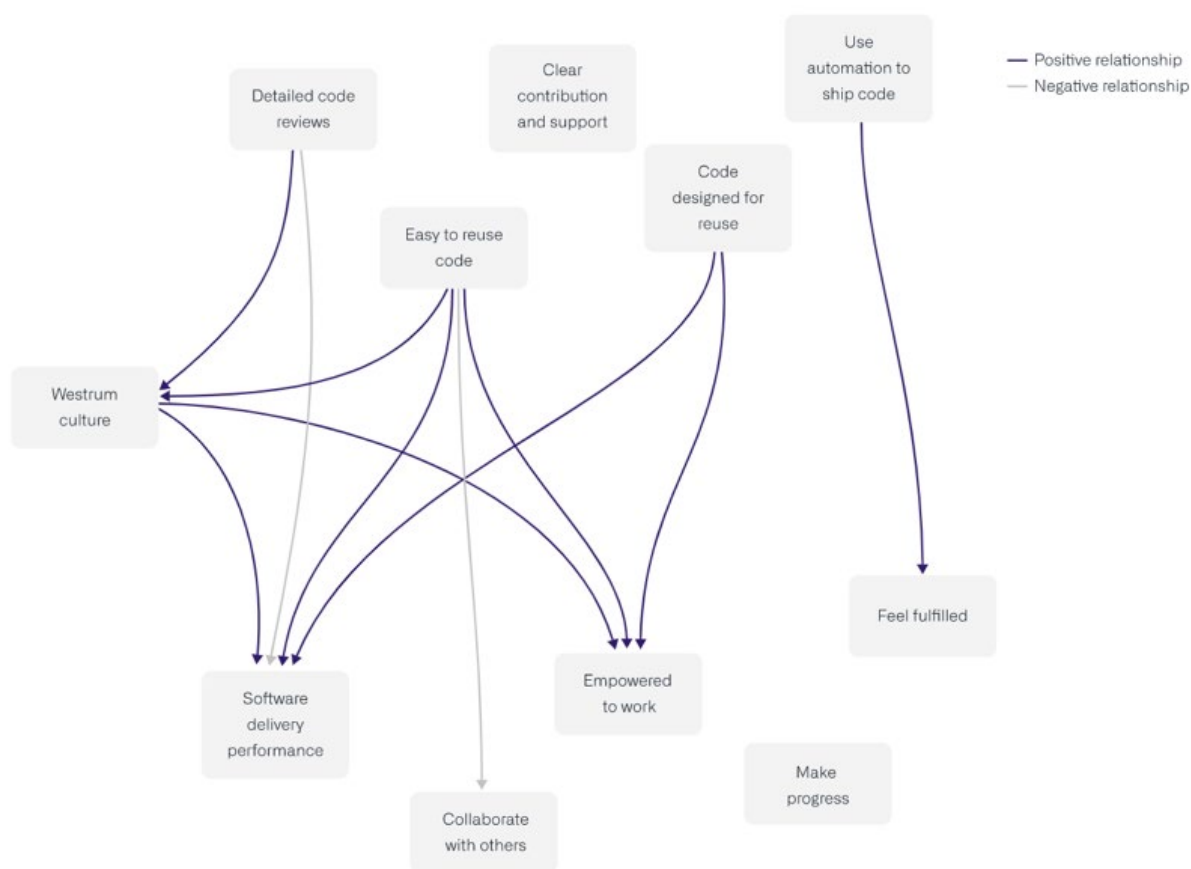
## 2.2 开源模式商业优势

### 2.2.1 开源项目可以为软件的开发提供杠杆

开源带来了产品的快速迭代与用户场景的拓宽。开源模式下的协同开发平台使产品迭代速度加快。开源模式下，社区成员能够协同编写软件，软件用户可以提交及时使用反馈，在反馈的基础上开发人员可以进行更高效地修改、迭代。此外，开源使用户可以直接参与到产品开发中来，这比闭源软件开发更能了解客户的需求，并且发掘出用户需要的使用场景。因此，相对于单一的闭源软件供应商，开源社区有着

更清晰的研发方向，能致力于开发出更多的功能，并创造出对用户更有价值的产品，避免不必要的开发成本支出。除此之外，企业可以聘用贡献者来保证开发的活跃性，向开发者提供灵活的工作地点与模式，提高自身的人才储备来加速产品开发。

图表 5：开源提高开发人员生产力

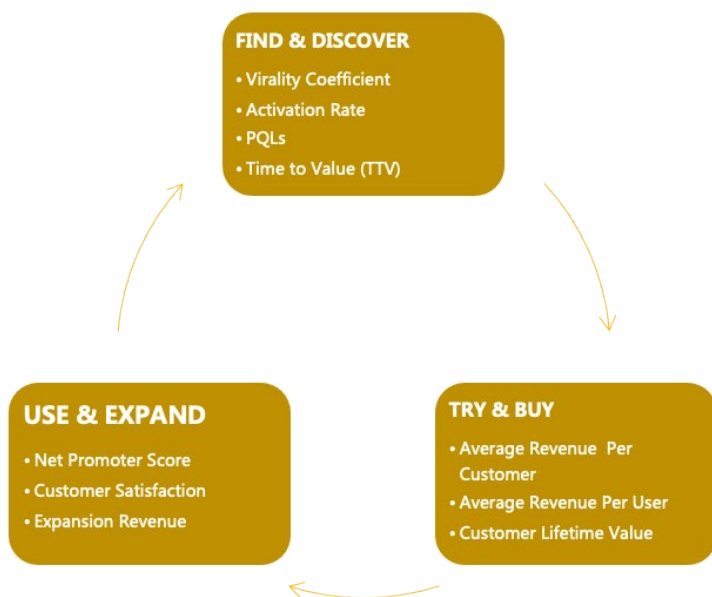


资料来源：GitHub

## 2.2.2 开源软件具有的 Product-Led Growth (PLG) 模型特征可以促进付费转化

PLG 模式主要通过自下而上的销售模式进行获客，产品是整个销售过程的核心。PLG 模式的增长飞轮有三个主要阶段：获客、转化、留存。在这三个阶段中，开源都有着区别于传统商业模式的优势。

图表 6：开源软件增长飞轮



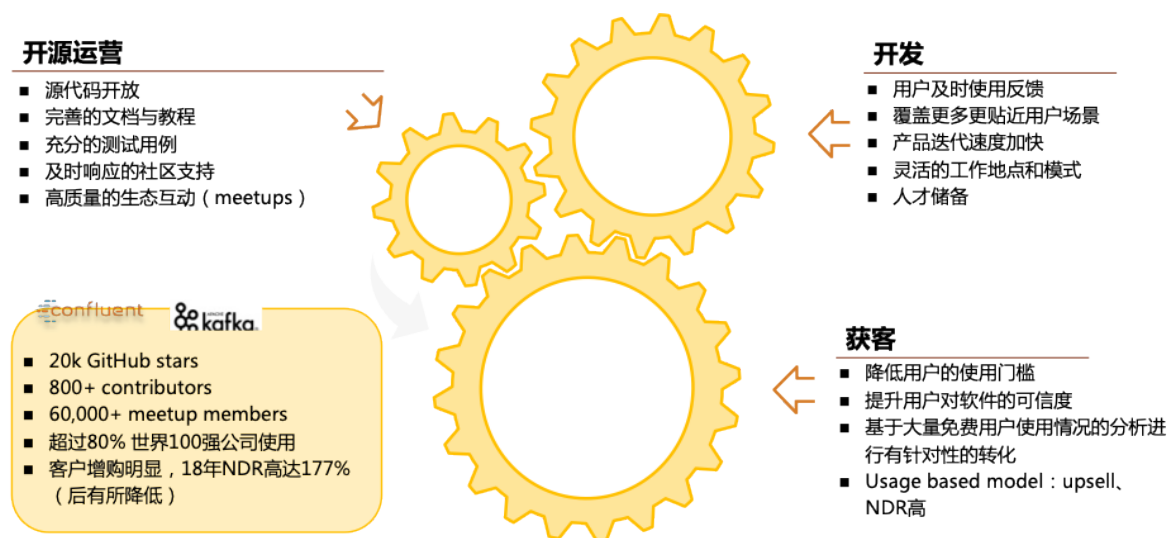
资料来源：云启资本

首先在获客阶段，开源运营模式降低了获客成本，并且使获客流程更具针对性。开发人员的相互交流、GitHub 等平台带来的社区型协作，加速了传播获客。开源产品的初始客户定位通常为开源社区的参与者，他们往往是企业里的开发者或者 IT 人员。培育了这些优质潜在客户，也就具备了“群众基础”。社区帮助打开企业的边界，让好的开源项目和产品的口碑传播得以可能。使用者为了解决自身问题和痛点自发的进行下载使用，此时开源软件产品不仅仅是作为通过功能解决用户问题的一个方式，也可以成为帮助企业去传播和增长的一个载体。长期来看，就可以降低企业的获客成本，让自动化的获客越来越多，降低销售方面的费用支出。

其次，在转化阶段，相比较传统商业软件，开源软件往往拥有更高的付费转化率。一方面，当用户使用过免费版的软件后，只要软件的功能可以很好的满足用户需求，就可以以较短周期的速度进行付费转化，并使其成为长期用户。另一方面，企业可以通过观察用户对免费版软件的使用行为，进行有针对性的转换跟进和追加销售，例如，向销售团队提供超出其使用限制并准备付款的客户列表。除了传统的销售转化，还可以通过自助购买路径进行转化（Self-service selling），这种转换路径很大程度上降低了销售成本。

最后在留存阶段，开源软件可以使用户规避供应商锁定风险，使其愿意进行长期使用。基于同一个开源项目，其下游可能会出现多个提供相似功能软件的供应商，并且可以以比较小的成本来改变供应商的选择，因此用户可以放心地选择长期使用软件。相反地，当顾客使用闭源产品时，如果在使用一段时间后想要转换使用另一个软件，需要重新进行硬件、数据等的部署，造成不小的转移成本，因此当用户选择使用闭源软件时，可能会由于软件后期开发情况不满足需求或者转移成本过高而放弃对软件的继续使用。

图表 7：开源项目可以为软件的开发和产品获客提供杠杆



资料来源：云启资本

### 2.2.3 开源可以带来技术和业务创新的良性循环

只有当技术创新与商业创新相结合时，开源的全部潜力才能实现。Andreessen Horowitz (a16z) 分析指出，开源是一种技术驱动的模式，它可以加速产品反馈和创新、提高软件可靠性、扩展支持、推动采用并汇集技术人才。然而正是有支持付费、Open Core 和 SaaS 模式等商业模式的出现，才有了现在的开源发展。经济利益创造了一个良性循环。开源的业务创新越多，开发者社区就越大，这会刺激更多的技术创新，从而增加对开源的经济激励，形成一个良性循环。

开源 3.0 时代下的技术创新与商业创新正在齐头并进。技术层面上，人工智能、开源数据和区块链是新兴创新的一些例子。而得益于时刻保持创新的开发者社区，开源已成为云计算，SaaS 服务，下一代数据库，移动设备，互联网甚至区块链的基础。下一代商业模式可能包括广告支持的 OSS，类似于大型专有企业支持的开源项目、数据驱动的收入以及通过区块链货币化的加密货币。

图表 8：开源的良性循环



资料来源：a16z

图表 9：开源 3.0 技术与商业模式的创新

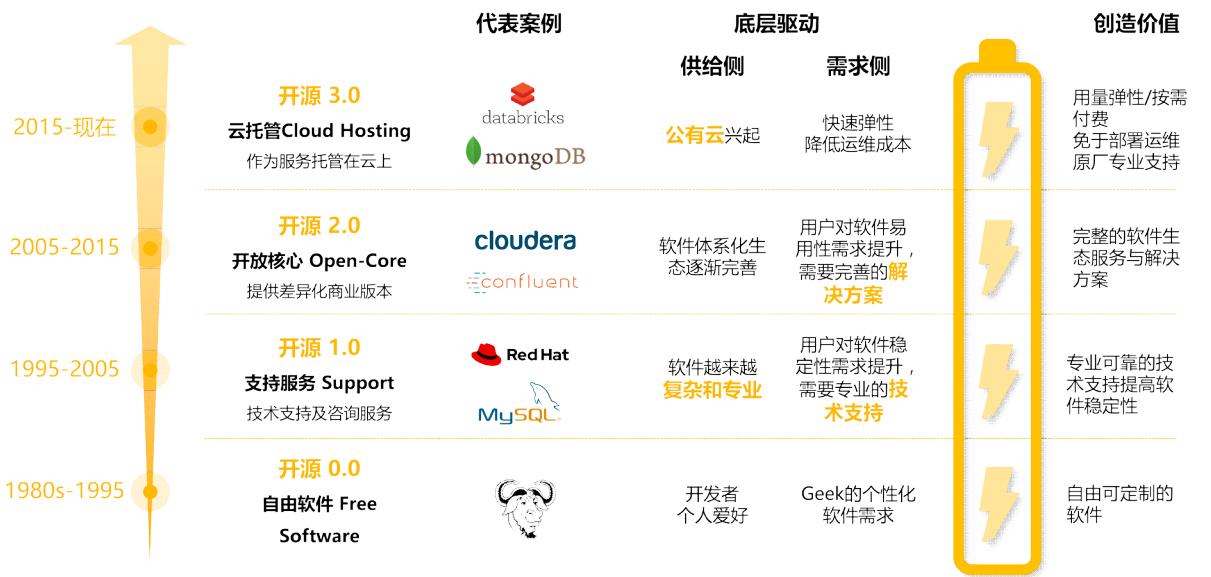


资料来源：a16z

2.3 开源软件商业化模式发展历程

开源软件商业化发展经过多次变革，从最初始的自由软件，逐渐发展出不同的商业化模式——从以 RedHat 为例的支持服务时代演变到当前的云托管时代。商业模式的发展演变离不开每个时代下供给侧与需求侧的驱动，在技术与经济大环境下的不断推动，开源行业才得以如今的蓬勃发展。

图表 10：开源商业化发展历程



资料来源：云启资本

开源 0.0 时代

开源起源于“自由软件”时代。最初是为了对抗大型闭源互联网公司对技术的垄断，开发者自发地开发了可以自由使用的“自由软件”。自由软件运动的发起人，也是 GNU 操作系统开发人及自由软件基金会创始人的 Richard Stallman 将自由软件定义为“用户可以自由地运行、复制、分发、研究、更改和改进的软件”。在这个时期出现了首批开源许可证 GNU（通用许可证，现在通常称为“GPL”）和 BSD 许可证。

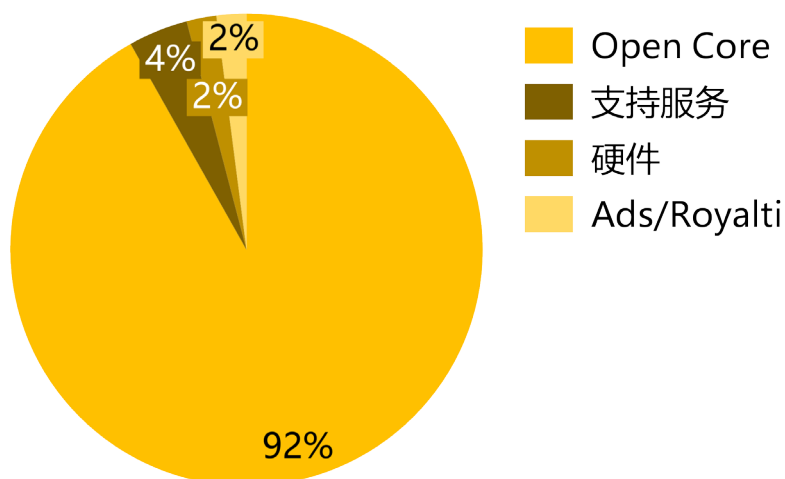
## 开源 1.0 时代

“Support 支持服务”模式面向开源软件客户。随着越来越多基础开源技术的出现，软件的复杂性和专业性都大幅度提高，用户对软件稳定性的需求也同步提升，需要专业的技术支持。这个时期出现了一些企业开始尝试基于开源软件实现商业化运营，主要的商业模式为“Support 支持服务”模式，为使用开源软件的客户提供付费的技术支持与咨询服务。

## 开源 2.0 时代

“Open Core 开放核心”模式成为主要的开源商业模式。随着越来越多的企业开发者加入到开源生态中，软件体系化生态逐渐完善，用户对软件易用性需求提升，此时用户更需要完善的解决方案。因此出现了新的收入模式，供应商将专有部分打包成与开源基础部分连接的单独模块或服务，或者在专门的商业版本中发布。用户可以通过付费的有更多功能模块组成的商业版本软件来获得更完整的软件生态服务与解决方案。Open Core 开放核心模式是当前成功的商业开源软件公司使用的主要商业模式，在年收入超过 100 万美元的公司中，有 92% 的公司都选择了 Open Core 的商业模式。

图表 11: 年收入 \$100M+ 的开源软件商业公司商业模式统计



资料来源: Crunchbase, 公开资料, 云启资本

## 开源 3.0 时代

云计算打开了 SaaS 服务模式的大门，开源软件目前更多的作为服务托管在云上。自云计算技术发展以来，云增长持续超出预期。不断增长的对灵活和可扩展基础设施的需求推动了 IT 企业的云计算支出与全球范围内云渗透率的不断提高。在这样的技术背景下，用户对降低软件运维成本的需求不断增加。一些开源软件公司例如 Databricks、HashiCorp 等提出了新的解决方案，通过 SaaS 使客户跳过内部部署，

直接将软件作为服务托管在云平台上。客户通过订阅 SaaS 服务，将前期高额的资本性支出转为小额的经常性支出，并在很大程度上缓解了运维压力。

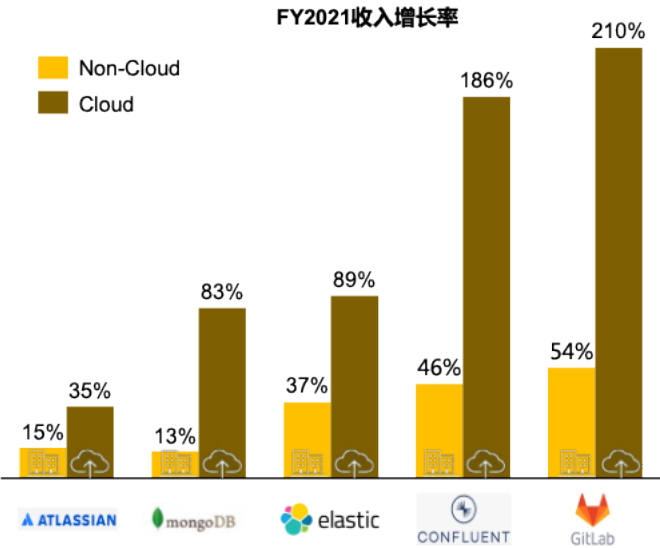
图表 12: 云计算支出增长以及全球云渗透率的提高持续受到推动

|                                            | 2015   | →    | 2020   | →    | 2025E  |
|--------------------------------------------|--------|------|--------|------|--------|
| 云计算支出                                      | \$105B | 3.0x | \$320B | 2.7x | \$848B |
| 渗透率<br>(占全球 IT 支出的百分比)                     | 3%     | 2.7x | 8%     | 2.0x | 16%    |
| Cloud Titan 收入<br>(Microsoft, AWS, Google) | \$11B  | 7.7x | \$83B  | 3.6x | \$303B |

资料来源: Battery Ventures, 云启资本

开源软件企业着重加强云战略布局，以加速和推动持续性增长。开源软件企业的 SaaS 服务模块的营收正在成指数倍增长。例如，GitLab 在其最近提交的 S-1 文件中指出，类似于 SaaS 的全托管云产品在 2020 财年至 2021 财年间的 ARR 占比从 9% 增长至 16%，SaaS 业务收入同比增长 210%。Confluent 云收入年增长率为 186%，MongoDB 的云收入现在占其总收入一半以上，2021 年 SaaS 收入增长达到 83%。

图表 13 : 2021 财年部分开源软件企业云服务和非云服务的收入增长率



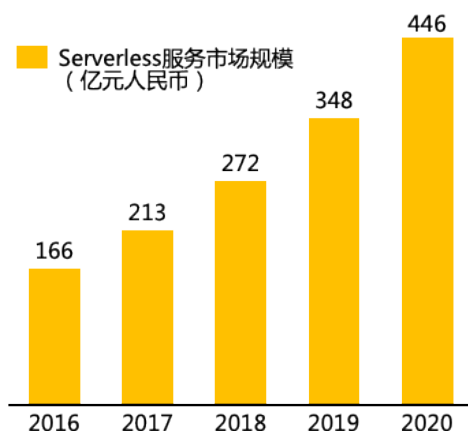
资料来源: 各公司公告, 云启资本

## 未来开源 4.0 时代

Serverless 将成为开源与云计算结合的新模式。无服务器（Serverless）是一种云原生开发模型，可使开发者在无需管理服务器的情况下专注构建和运行应用。无服务器方案将服务器从应用开发中抽离，由云供应商负责置备、维护和扩展服务器基本架构，而开发者则可仅将代码打包到容器中进行部署。Serverless 的全球规模保持着比较高的增速。据沙利文报告统计，2020 年营收统计口径下，Serverless 全球市场规模高达 446.1 亿元，相对于 2019 年 348 亿的市场规模，增速达到 28%。

通过在自有云和 Serverless 基础上运行开源，开源方案成本降低。Serverless 平台会根据请求的数量来创建对应的函数实例来执行，无需人工干预，瞬间弹性扩容，应对流量爆发。更大程度上降低了云基础设施成本以及维护扩张成本。除此之外，基于 Serverless 开源解决方案能够以零支出在世界各地部署，并且在运营成本和管理复杂度方面上不会有任何增长。同时，还能够提供更优异的终端用户性能。

图表 14：2016–2020 年全球 Serverless 服务市场规模（按营收计）



资料来源：沙利文，云启资本

## 3 开源软件公司商业化路径选择

开源项目的开发者应该意识到开源不是免费的，开源可以进行商业化。从开源这一术语的出现至今，众多的个人开发者以及企业在开源商业化的道路上不断探索，而历史的开源商业化演变道路也向我们展示着开源项目可以成功的进行商业化。这是由于开源这一独特的产品开发模式产生了自己的商业价值，开发者可以选择不同的商业模式将产生的商业价值实现货币化。

## 3.1 开源软件商业化可以选择的路径和案例

**Support 支持服务**——为免费的开源软件提供付费的技术支持和咨询服务，例如集成、使用培训、商业部署等。支持服务的优势在于可以与客户深度融合。但是这种模式从长远来看会受到限制，主要原因为：(1) 支持通常需要大量的人工工作，因此降低了业务利润；(2) 工作缺乏可复制性，部署 / 集成不可扩展，无法规模性发展；(3) 客户转化率低，通常只能将不到 1% 的用户转化为付费客户，主要是因为只有那些依赖关键任务系统项目的客户才愿意为支持付费。然而，那些严重依赖项目的人自然会随着时间的推移投入自己的工程努力来了解项目，从而减少对外部支持的需求。选择 Support 支持服务商业路径的主要公司案例是 Red Hat。

**Hosting 托管**——供应商将其开源软件作为服务托管在云上，通过收取每月 / 每年的托管和服务费获利。托管意味着提供一个完全托管的项目版本，当用户想要在生产中部署时，可以使用该软件启动远程服务器，而不必担心它不用担心备份、停机、升级等问题。Databricks、Acquia 等公司已经成功证明了托管模式商业化路径的可行性。但是这也引起了公有云提供商（例如 AWS）与开源社区的矛盾。

**Restrictive Licensing 限制性许可**——通过提供一个带有稍带限制的开源许可证来激励使用者进行付费。GPL 和 AGPL 许可，以及新创建的 Commons Clause（被某些 Redis 模块采用）都是这种模型的例子。特别是 AGPL 和 Commons Clause（以及 MongoDB 推出的新 SSPL）也是旨在防御公共云提供商的许可证。这种方法的最大缺点是这些许可证会影响软件采用，通常会流失潜在用户。尤其是部分大公司禁止使用该模式下的开源软件。

**Open-core 开放核心**——供应商将专有部分打包成与开源基础部分连接的单独模块或服务，或者在专门的商业版本中发布。开放核心模式是指公司的大部分代码库是开源的，而一小部分（针对生产或企业用户）是专有的。通常，专有功能是生产部署或大规模生产所需的功能。（例如，对于开源数据库，监控、管理、备份 / 恢复和集群等功能通常是专有的）。开放核心模式通过允许开源公司在专有代码库中保留某些功能以此来抵御公共云厂商的竞争。但是，这种模式存在两个挑战。首先，开源范围的尺度难以拿捏。如果开源部分太多，专有功能的盈利就会变少；但如果开源提供的功能太少，那么开源项目可能无法得到广泛采用。另一个挑战是，将开源与代码中的专有功能彻底分开有时比较困难。

**Hybrid Licensing 开放核心 + 混合许可**——混合许可在同一个代码库中混合了开源代码和专有代码。用户可以选择只使用开源代码，或者同时使用开源代码和专有软件代码。这种模式是基于开放核心模式的改进，因此优势包括 open-core 的所有优势，以及更多：(1) 将所有内容都放在同一个代码库中，可以更轻松地管理工程流程和开发；(2) 允许用户方便从免费升级到付费；(3) 允许外部社区成员（例如，通过 GitHub 评论、提交问题）为专有功能做出贡献。

图表 15：开源的五种商业模式

| 商业模式                               | 简介                                                                                                                                                | 特点                                                                                                                                       | 代表企业                                                                                                                                                                    |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Support 支持服务</b>                | <ul style="list-style-type: none"> <li>用户只需为技术支持及咨询服务买单</li> </ul>                                                                                | <ul style="list-style-type: none"> <li>人工外包作，利润率偏低</li> <li>工作可复制性低，scale较难</li> <li>客户转换率低，通常&lt;1%</li> </ul>                          |   |
| <b>Hosting 托管</b>                  | <ul style="list-style-type: none"> <li>供应商将其开源软件作为服务托管在云上，通过收取每月/每年的托管和服务费获利</li> </ul>                                                           | <ul style="list-style-type: none"> <li>该模式成为了部分云厂商打包开源项目赚取利润的途径</li> </ul>                                                               |   |
| <b>Restrictive Licensing 限制性许可</b> | <ul style="list-style-type: none"> <li>通过提供一个带有稍带限制的开源许可证来激励使用者进行付费</li> </ul>                                                                    | <ul style="list-style-type: none"> <li>许可证定义模糊，需要法院判决</li> <li>部分公司禁止使用该商业模式下的开源软件</li> </ul>                                            |   |
| <b>Open-core 开放核心</b>              | <ul style="list-style-type: none"> <li>该模式下的大部分代码是开源的，而少数代码（针对企业用户）是专有的，需要收费</li> <li>专有部分可以打包成与开源基础部分连接的单独模块或服务，或者在分支版本中分发</li> </ul>            | <ul style="list-style-type: none"> <li>该模式可以避免云厂商打包开源项目赚取利润</li> <li>难以拿捏开源范围的尺度</li> <li>很难将代码中的开源与专有特性完全分离</li> </ul>                  |   |
| <b>Hybrid Licensing 开放核心+混合许可</b>  | <ul style="list-style-type: none"> <li>最新的模式，在开放核心基础上进行了改进</li> <li>混合许可可在同一个代码库中混合了开源代码和专有代码</li> <li>用户可以选择只使用开源代码，或者同时使用开源代码和专有软件代码</li> </ul> | <ul style="list-style-type: none"> <li>代码在同一个代码库中，使管理和开发变得更容易</li> <li>允许用户方便升级到付费模式</li> <li>允许外部社区（比如GitHub）成员对专有软件功能模块进行改进</li> </ul> |   |

资料来源：云启资本

## 3.2 开源项目的 life cycle 以及软件供应商在对应阶段的重点工作

### 3.2.1 早期阶段：软件项目代码公开化及社区构建—开发者社区管理

在项目的早期阶段，供应商创建和管理开发者社区，在欢迎开发者加入成为社区成员的同时保持项目主导地位。个人、团体或企业为解决问题创建软件，并将软件项目代码公开化，这代表着一个开源项目的开始。供应商通过构建沟通渠道、技术平台和治理模型等基础设施来创建开发者社区。在社区创建时，供应商应该向开发者解释清楚项目的管理方式、项目主管、外界可作出哪些贡献、以及项目是否由第三方基金会（如云原生计算基金会 -CNCF 和 Apache 软件基金会）监督。除此之外，在这个阶段企业应该进行概念验证（PoC）证实开源项目的行业价值与机会，主要可以通过注册用户以及下载量来作为衡量指标。

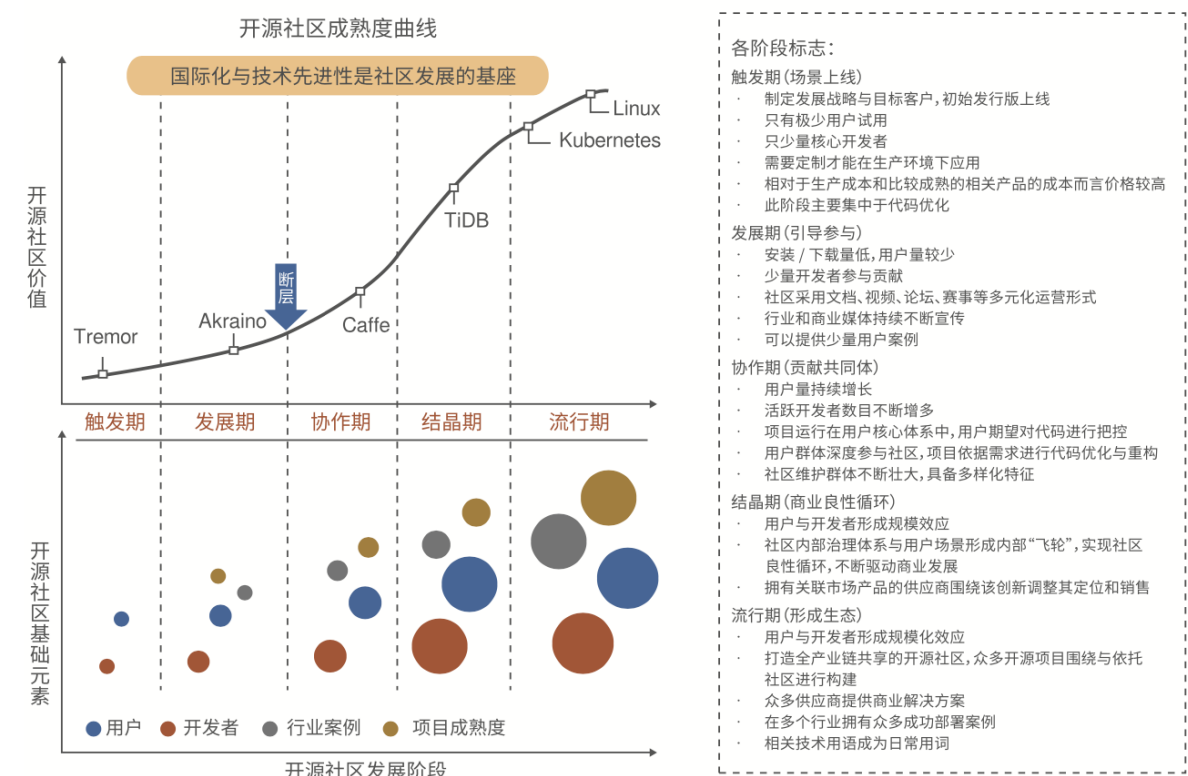
### 3.2.2 成长阶段：Roadmap 策略制定及扩大社区建设—产品管理

在开发者社区进行有效构建后，下一个阶段是指定产品路线并加速扩大社区。企业应明确软件产品的专有和开源路线图，分析产品特点及使用信息反馈以预测销售机会。目前大多数开源软件企业采用 Open Core 开放核心模式以及 SaaS 托管模式，旗下产品或服务通常分为免费版以及付费版，因此开源企业通常需要管理开源产品路线图和专有产品路线图这两个路线图，如何确定哪些功能开源，哪些功能专有是产品管理中非常重要的部分。

当一个开源项目刚刚开始的时候，需要重点吸引新开发者用户，建立项目“品牌”。企业应该持续投入反馈社区，打造高质量产品及内容的产出，以此维护社区信任，吸引更多的开发者加入社区来加速社区扩张。

在社区运营方面,项目创始人往往是社区的灵魂人物。创始人可通过参加面向开发者的会议与在线讨论会,并撰写有关该项目的技术文章、参与论坛等等方式进行项目的宣传。要注意的是,在此阶段可能会发生社区成熟度落后于宣传热度而导致的断层,开源项目发展路径和用户需求出现偏差,或者优化迭代速度无法满足用户要求,造成用户流失,社区死亡。因此,在这个阶段对项目 and 社区的大量投入是非常必要的,如果可以跨越这一阶段,会迎来快速增长时期。

图表 16 : 开源社区成熟度曲线



资料来源: PingCAP, 中国信通院

3.2.3 加速扩张阶段：加速软件采用及付费转换—销售管理

加速扩张阶段的主要任务是增强开发者和用户对软件的喜爱、采用和价值,发掘潜在客户并将免费用户向付费用户转换。定位潜在客户应该优先考虑针对特定细分市场的活动,基于产品使用信息进行用户分析,了解哪些角色和部门正在使用该产品以及他们的兴趣所在,同时预测免费用户向付费用户转换的百分比。

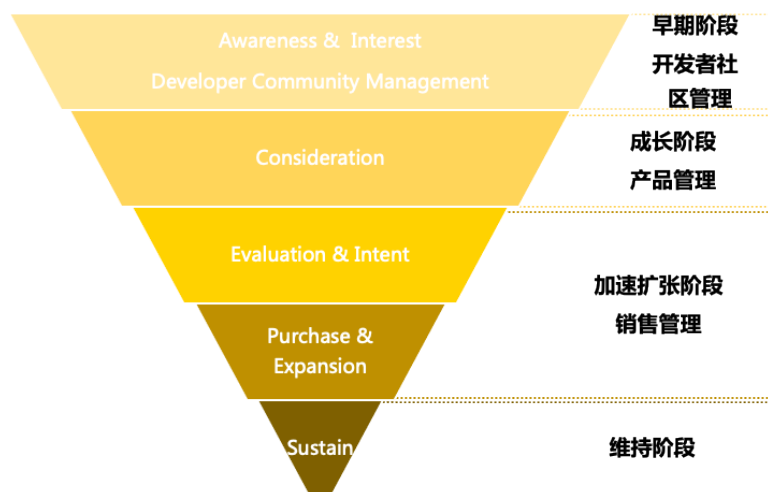
开源软件企业在进行付费转换时要遵循 PLG 增长飞轮的获客、转化、留存三个阶段,分别采取不同的行动来增大转换率。获客阶段,企业应着重关注产品合格线索 (Product Qualified Leads – PQL), PQL 是根据产品的使用行为,去把用户确定为一个潜在的销售线索。获客阶段可以参考的量化指标有 PQLs、

TTV（新用户到达他们的“aha moment”或激活事件并实现价值所需的时间）等。在转化阶段，企业应该根据免费版软件的使用行为反馈，进行有针对性的转换跟进和追加销售。最后在留存阶段，产品质量是能否留住客户的最根本因素。只有持续地为客户提供价值，才能提高客户留存率。开源软件的销售也可以兼容传统 SLG 模式和 PLG 模式。一方面，PLG 模式免费版产品面向个人，或中小企业，或者大企业员工用户，扩大用户基数；另一方面，通过 SLG 销售模式去面向中大型企业进行大单交易。

### 3.2.4 成熟阶段：维持开源社区运行

在项目成熟阶段，供应商主要任务是发布新版本、提供漏洞修复、维护社区的运行。开源社区是开源项目的根基。自由包容、积极活跃的开源社区支持项目持续稳定的发展与迭代。

图表 17：开源项目生命周期



资料来源：云启资本

## 3.3 商业化过程中的风险点

### 3.3.1 开源项目技术归属风险

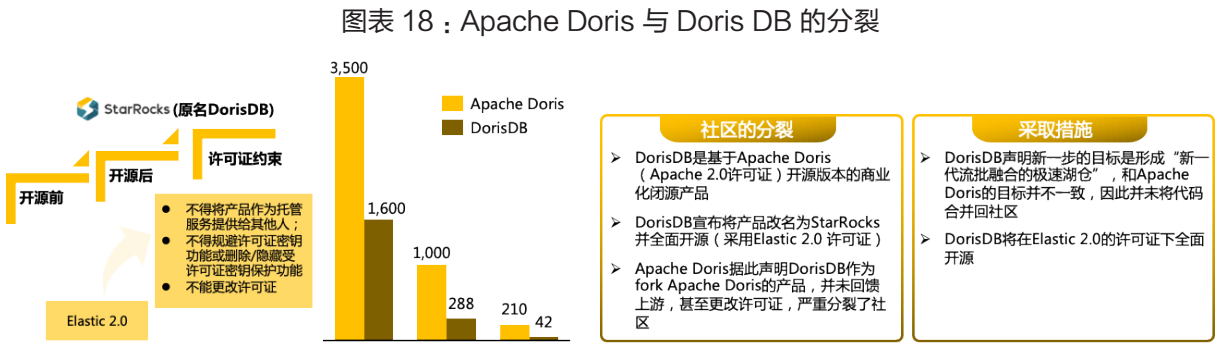
开源项目代码公开透明，可能会存在多个下游供应商基于上游社区技术进行各自的商业化开发，这会导致开源项目技术归属出现争议。Apache Doris 和 StarRocks（原 DorisDB）就存在由于开源技术归属引发的矛盾

#### 专家点评

姜宁：这个问题还是集中在商标权的问题，不是技术归属的问题。开源许可协议允许在满足一定条件下的二次分发，由于 DorisDB 和 Apache Doris 有很强的关联性，存在误导大家的嫌疑，需要及时纠正。

Apache Doris 2018 年由百度贡献给 Apache 软件基金会。2020 年 2 月，百度 Doris 团队的个别成员离职创业，基于 Apache Doris 之前的版本做了自己的商业化闭源产品 DorisDB，也就是 StarRocks 的前身。DorisDB 在 2021 年 9 月宣布采用 Elastic 2.0 许可证下全面开源。这件事引起了很大争议。

Apache Doris 认为 DorisDB 作为 Fork 自 Apache Doris 的项目，不仅未回馈上游社区，选择变更为 OSI 不认可的 Elastic 2.0 许可证，增加了“不得将产品作为托管服务提供其他人”、“不得规避许可证密钥功能或删除 / 隐藏受许可证密钥保护的功能”、“不能更改许可证”等条件，是一种“伪开源”的行为。但是由于 Apache Doris 使用的许可证为 Apache License 2.0，允许分发完全自由、允许项目代码被修改、允许作为开源或商业化软件再次发布，法律层面上 Apache License 约束不了这样的行为。在这件事上，双方各执一词，引起了人们对开源项目技术归属问题的重视与思考。



资料来源：公开资料，云启资本

3.3.2 代码安全风险

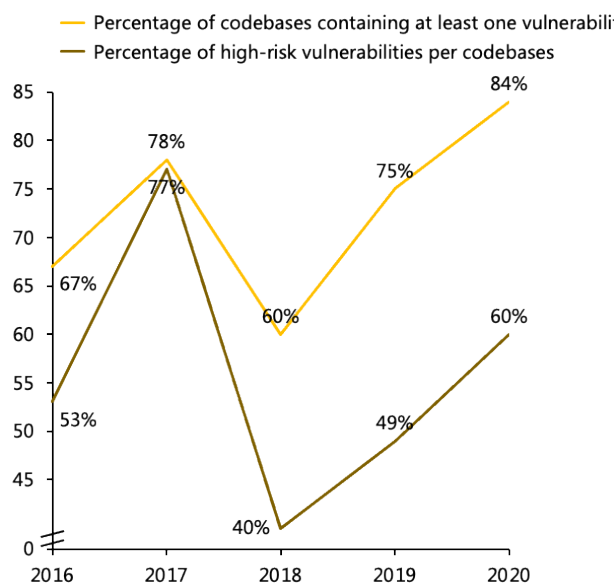
开源组件漏洞引起的安全风险也是开源风险中不可忽视的重要因素。开源代码具有公开易获取的特点，这为开发者提供便利的同时，也隐含着漏洞反馈和修复滞后的潜在风险。开源项目一旦被广泛使用，一方面漏洞信息散落在各类开发者手中，能否及时被官方收录是一个挑战；同时另一方面，如果软件使用者跟踪漏洞修复不及时，则其被攻击的风险将大大提升。

专家点评

姜宁：开源组件的使用并不是没有成本的。一般来说开源软件许可中都包含了一个不提供质保的声明。如果想要比较好解决开源项目代码安全风险的问题，公司要么投入人力参与上游项目的开发，要么购买商业公司基于开源项目制作的发行版。

大量开源项目、开源组件存在安全漏洞，且漏洞数量近年来逐年递增。据 Synopsys 统计，在 2020 年审计的 1,500 多个代码库中，其中 84% 至少包含一个公共开源漏洞——比 2019 年的 75% 增加了 9%。而包含高风险开源漏洞的代码库在 2020 年增加到 60%，比 2019 年审计的 49% 增加了 11%。

图表 19：开源代码漏洞统计



资料来源：Synopsys，云启资本

全球知名开源日志组件 Apache Log4j 于 2021 年 12 月被曝存在严重高风险级别远程代码执行漏洞，引发人们对开源安全性问题的探讨。12 月 9 日，Apache Log4j2 被曝出第一个高危漏洞 Log4Shell，并在此之后持续爆雷，至 12 月 22 日已经发现了第三个高危漏洞 CVE-2021-45105。而由于 Log4j 在国际上的流行度，漏洞带来的安全问题是巨大的。根据谷歌安全团队的统计，截至 2021 年 12 月 16 日，来自 Maven Central 的 35,863 个可用 Java 组件依赖于 Log4j。这意味着 Maven Central 上超过 8% 的软件包里至少有一个版本会受此漏洞影响。除此之外，根据 Cloudflare 研究人员的观测，每秒有超过 1,000 次利用 Log4j 漏洞的尝试。有不法分子利用远程代码执行漏洞窃取云基础设施，部署加密货币矿工和勒索软件。据 Check Point 的统计表示，全球近一半企业因为该漏洞受到了黑客的试图攻击。

图表 20：受 Log4j 漏洞影响的企业网络的区域统计

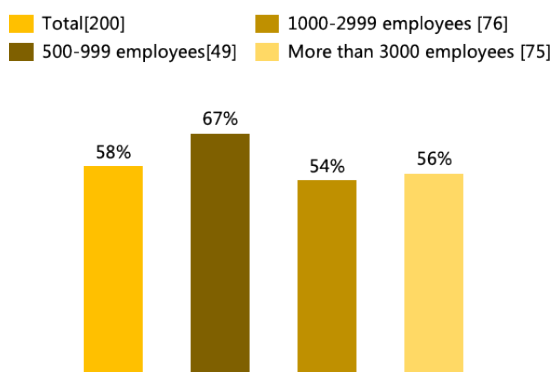


资料来源：Check Point Research，云启资本

### 3.3.3 云厂商竞争风险

与使用开源项目但不回馈的公共云公司的竞争是多数软件供应商面临的挑战。云厂商通过 Fork 流行项目的源代码，将其作为付费托管服务面向客户，用基于原供应商代码的项目与其竞争客户，造成了云厂商与开源项目现在的摩擦现状。在 VansonBourne 的 2021 年开源调研报告中显示，对于 200 家开源公司进行问卷调查，58% 的公司表明其正在面临着与使用开源项目但不回馈的公共云公司的竞争。作为主要开源项目维护者的公司对于云厂商非常抵触，因为他们投入大量资源为开源项目做出巨大贡献，这些项目作为他们自己的专有付费产品的基础。但是使用开源项目但不回馈社区的云厂商，削弱他们的产品并冲击他们的市场。

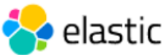
图表 21：开源公司是否正在面临着与使用开源项目但不回馈社区的云厂商间竞争



资料来源：VansonBourne，云启资本

围绕这个问题以亚马逊网络服务 (AWS) 与 Elastic 的争论为例。Elasticsearch 是 Elastic 公司旗下的一款搜索和数据分析引擎，采用 OpenCore 的商业模式。AWS 在它的下游开了一个发行版叫 OpenDistro for Elasticsearch。2021 年 1 月，Elastic 公司宣布即将变更 Elasticsearch 的 Apache License 2.0 为双授权许可，即 Server Side Public License (SSPL) + Elastic License，主要是针对 AWS 基于 Elasticsearch 的托管服务。作为应对措施，AWS 宣布创建“真正”开源的 Elasticsearch 分支，将 Elastic search 修改协议之前的最后一个版本 7.10 直接 fork 出来变成 OpenSearch，开源许可证也会继续使用 Apache License 2.0。业界对于 Elastic 改变许可证的做法褒贬不一，有些人认为是在合理的保护自己对于软件的贡献成果，有些人认为 Elastic 已经违背了开源的原则，实际上已经走向了闭源。

图表 22 : Elastic 和 AWS 之争

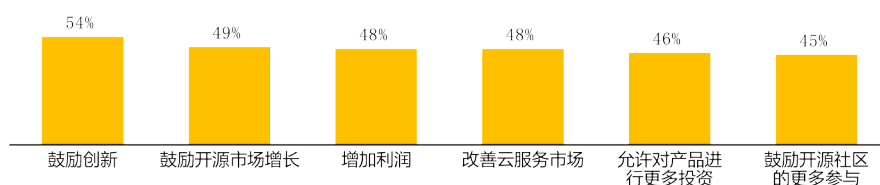
| 开源项目                                                                                               | 供应商                                                                                                                                                                                          | 修改前                                                                               | 修改后                                                                                | 修改后的变化                                                                                                  |
|----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| <br>elasticsearch | <br>elastic<br><br><br>aws |  |  | SSPL要求，如果将开源项目作为服务提供，需要将Service Source Code开源（所有用于项目提供为服务的代码，包括管理软件、UI、API、自动化软件、监控软件、支持软件、存储软件、托管软件等）。 |

资料来源：公开资料，云启资本

软件供应商通过改变许可证限制程度作为抵抗云服务商的做法存在两面性。为了防止云供应商产生的威胁，大多数开源公司选择了改变许可证限制条件的方法。例如 Redis 创造了新的许可证 Redis Source Available License (RSAL)。大型开源公司 MongoDB、Confluent 和 Elastic 也转向了自己的定制许可证。然而，这些变化对开源社区以及供应商来说仍然存在争议。

对开源软件企业进行访问时，一些企业表达了更改许可证限制对于企业可能存在的好处。例如当开源公司可以将成果保护起来时，可以激励其在开发时投入更多的资源。还有其他的一些优点，如鼓励创鼓励开源市场增长、增加利润等等。

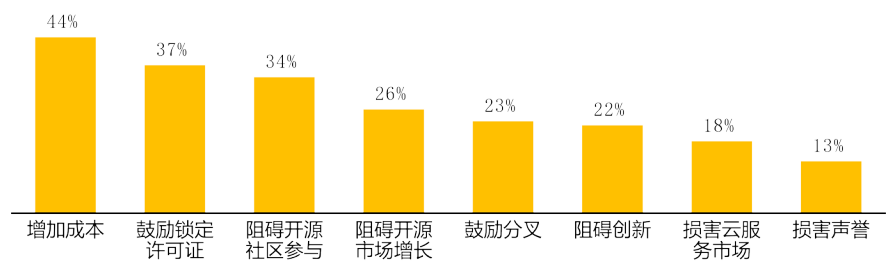
图表 23 : 更改许可证对开源软件公司的优势



资料来源：VansonBourne，云启资本

尽管更改代码许可有助于抵制公共云公司，但这种做法将会给那些确实想要贡献并按预期使用开源的人带来不好的影响。不可避免地，开源公司转向可用的源许可证会增加成本并鼓励锁定许可证，这违背了开源的本意—开放。除此之外，更改许可证也可能带来鼓励分叉、阻碍创新、损害云服务市场等弊端。

图表 24：更改许可证对开源软件公司的劣势



资料来源：VansonBourne，云启资本

开源软件商应该加强自身产品差异化的能力，主动拥抱云服务商进行合作。在开源软件公司与云厂商的竞争中代码不是竞争护城河，而是社区以及开发。独立开源软件公司具有三大竞争优势：企业客户不希望供应商锁定；用户更想从编写代码的人那里购买软件及服务；云厂商没有开源公司的专长。首先企业客户不希望供应商锁定。应用云厂商的托管服务，要将其所有对应数据托管到其平台上，可能存在比较大的转移成本，造成供应商锁定。在这一方面开源可以避免供应商锁定，更受企业用户的喜爱。其次，软件用户更倾向于向编写代码的公司去购买产品。作为开源社区的创建者与主导者，软件用户对开源软件公司的信任度更高，更愿意从开源软件公司购买软件而非云厂商。再次，开源软件公司有着自己的专长。开源软件公司可以吸引大量的开发者来完善软件和服务，这些开发者往往具有云厂商团队所不具有的许多技能。

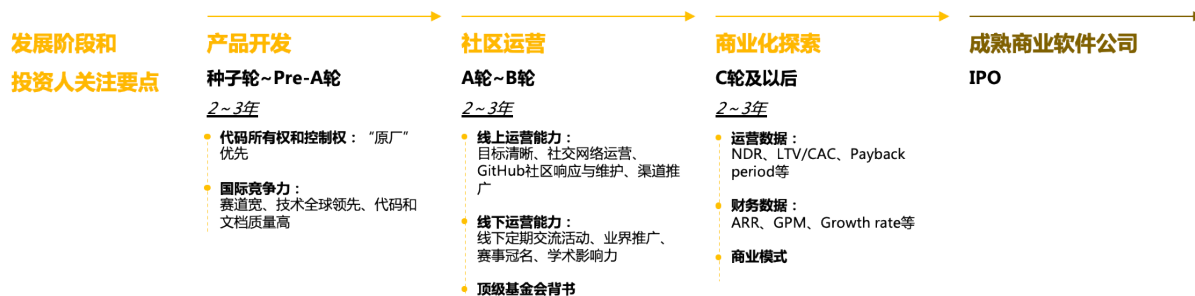
在实际情况中，云和开源并不是“零和博弈”，而可以是相互促进的合作关系。云是开源软件新的发展机遇。开源产品本身的特点使其十分适合在云上运行，云帮助这些开源软件构建起了新的商业模式。云计算可以解决基础软件之间、企业之间存在太多差异化的问题，开源软件的优质功能与云计算相结合可以高效的实现商业价值。PingCAP 创始人黄东旭也曾表达了自己的观点—云是开源可规模化变现的唯一出路。他指出一个好的生意应该是可以规模化的，但是传统开源商业模式需要人的介入（销售 / 售前 / 售后交付等等），而基于人的生意是没法规模化的。云的本质是一个资源租赁（Hosting），用户按需购买，让开源的规模化成为可能。

开源软件商应该加强自身产品差异化的能力，若开源产品做的足够复杂，云服务商就不能进行简单的打包销售，必须要同开源软件商进行合作。云厂商可以进行底层整套生态的培育。云、开源、社区之间能够形成生态，从而产生正面影响。

## 4 投资——如何找到下一个开源独角兽

### 4.1 成功商业开源项目判断标准

图表 25：不同阶段投资人对开源项目的判断要点



资料来源：云启资本

#### 4.1.1 拥有代码所有权和控制权——产品开发阶段

拥有代码所有权和控制权意味着拥有项目主导权、开发自主权、知识产权和商业模式选择权。一个开源软件如果要成功地商业化，最好背后只有一家商业化公司。这也是一个历史教训，Cloudera 和 Hortonworks 都是基于 Hadoop 的上市公司，但是因为基于同一个开源项目，中间有很多斗争内耗，而且由于产品没有明显的差异化，最后价格战的结果就是两败俱伤。相反的，正面例子既有 2021 年上市的基于 Kafka 项目的 Confluent，还有基于 Spark 的商业化公司 Databricks，他们的特点是只有一个商业化公司在主导项目的产品和商业化进程，不会有内耗、价格战的问题。

图表 26：原厂开源项目 VS 非原厂开源项目



资料来源：云启资本

#### 4.1.2 具备国际竞争力——产品开发阶段

开源项目的国际竞争力主要通过赛道选择、技术水平、获客能力、社区运营能力来体现。开源本身作为

基础软件来说是没有国界的，全世界的诉求一致。要脱颖而出的话必须做到同领域的世界领先地位，而不仅仅在中国内领先。由于研发人员可以采用开源软件时可以选择美国的也可以用中国的，不会牵扯到代码安全或者自主可控的问题。因此如果开源项目没有全球竞争力，那它的市场就会特别局限。以 PingCAP 为例，它的数据库产品 TiDB 对标的是美国的 CockroachDB，但是无论从社区活跃度、GitHub 的星标或者一些评论文章，PingCAP 的指标已经全面超过了它的美国对标。

图表 27：如何提高开源项目国际竞争力



资料来源：云启资本

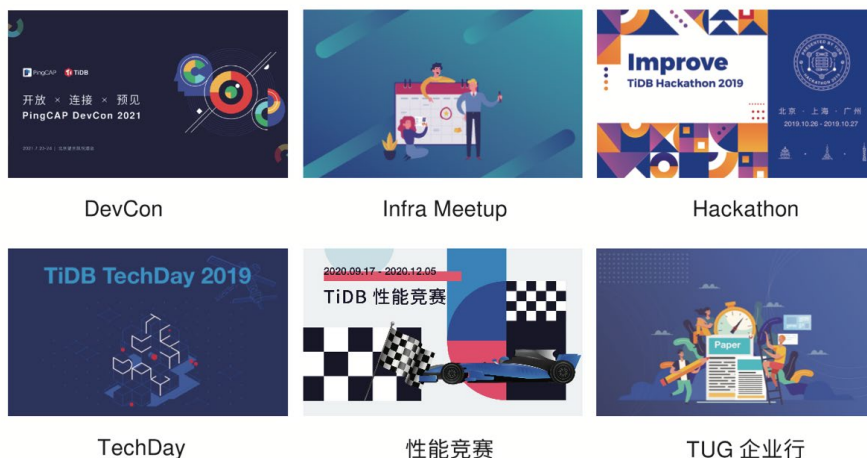
4.1.3 社区运营能力——社区运营阶段

开放、协作的社区是开源项目的第一支柱。一个成功的开源项目一定有一个活跃的开源社区，不论是参与开发的开发人员还是参与使用的用户。评价指标可以用包括 GitHub Star、Fork、Pull Request、Contributor 等来衡量。从 GitHub Star、Fork 等可以看项目参与的广泛程度，从 Contributor、Commits 可以看出项目的迭代情况，这些都是比较量化的指标。而从问答则可以看出社区的活跃度和了解真实用户反馈。尽管 OSS 社区的规模各不相同，但日益增长的人气是 OSS 项目激发开发人员群体浓厚兴趣的关键指标。

企业应具有强大的线上和线下宣传运营能够吸引更多的开发者参与到开源项目中，以此增加项目的影响力。线上运营方面，官网要有清晰的目标、蓝图和路线，让开发者理解项目本质，从而吸引开发者参与。并且要积极进行社交网络的运营，例如维护公众号等社交媒体，发表高质量的技术文章，以解决问题的心态帮助开发者等等。对 GitHub 社区应该进行及时的响应与维护，对开发者需求及时响应，定期维护

代码，及时更新说明文档。除此之外还应拓宽渠道推广，不断加强 SEO，增加项目线上曝光度和易得度。线下的运营方式包括：通过定期举行社区 meetup 等线下交流活动，增加开发者之间的交流，扩大社区的影响力；参加各种业界会议，并与媒体保持良好的合作关系，增加媒体曝光度；举办开源比赛，让全球的公司和学校参与进来，增加项目知名度；与学术界合作，发表顶级论文，增加业界影响力等。

图表 28：TiDB 社区运营模式

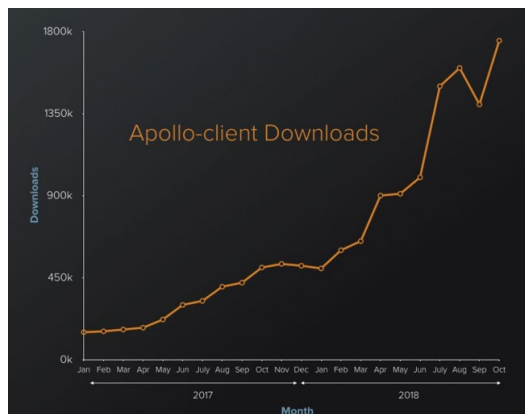


资料来源：PingCAP 官网

#### 4.1.4 市场匹配能力——商业化探索阶段

首先，开源项目应该具有产品 - 市场契合。产品与市场的契合体现在用户对软件的初始采用率，可以用下载量这一指标进行衡量。产品市场契合度主要是用来衡量免费产品的用户基数，因为在开源的商业模式下社区用户（含开发者）有一定几率在未来转化为付费客户，所以在初期可以吸引到多少用户进行产品的使用时后期付费转化的先决条件。

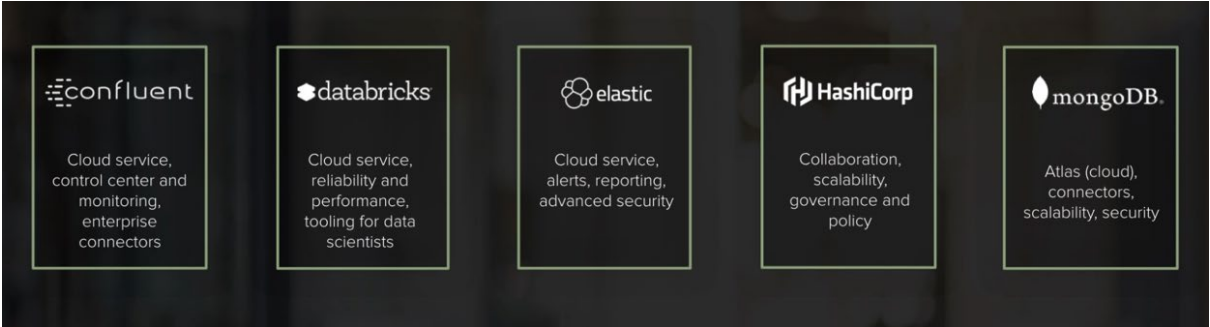
图表 29：Apollo 客户下载情况



资料来源：a16z

其次，开源项目应具有价值－市场契合。价值与市场的契合度就是要找到企业要进行商业化的内容来创造收入。价值与市场契合度是开源软件公司能否成功进行商业化最重要的因素。产品市场契合度引入了用户，价值市场契合度就是找到客户关心并愿意支付的费用，进而使企业可以自然延伸来推动收入。开源软件公司已经找到了一些围绕功能的价值市场契合，包括 RAS（可靠性、可用性、安全性）、工具附加组件、性能、审计、服务等。

图表 30：开源软件公司价值市场契合点



资料来源：a16z

#### 4.1.5 成熟的商业模式——商业化探索阶段

成功的开源公司必须有成熟的商业模式，选择哪种商业模式取决于可以为客户提供什么价值以及如何最好地提供这些价值。支持服务模式是开源 1.0 时代的基础模式，但是 RedHat 已经在这方面垄断了市场并实现了规模化。选择支持服务模式有可能要与 RedHat 在这方面竞争。Open Core 模型在开源软件之上分层增值专有代码，是内部部署软件的一个很好的模型。如果可以在不损害开源软件采用的情况下保持专有的高价值的组件，那么 Open Core 将是一个很好的模型。但是 Open Core 模式存在的风险是——如果不能很好地决定哪些功能属于专有，哪些属于开源码时，社区疏远可能会成为一个问题，会出现项目的分叉版本，或者围绕相同的代码库启动一个新项目。如果企业的价值和竞争优势在于软件的运维，那么可以选择 SaaS 作为商业模式。但是要面对公共云厂商竞争的潜在风险。

不是所有开源软件都适合做商业化。根据对美国开源独角兽公司的相应研究，发现大数据、AI、企业搜索、中间件和操作系统这几个领域相对是容易商业化的。相反一些前端的控件，关注度很高，但是很难设置收费点，商业化会有难度。因此成功的开源公司有成熟的商业模式，并且通常在美国有对标的项目。

### 4.2 中国开源市场

随着越来越多的中国优秀开发者参与到开源社区中，涌现了很多达到国际水平的优秀开源项目。在开源领域表现最为亮眼的是 PingCAP，TOP 20 项目中上榜的项目有 6 个，包括位于榜首的由其自主设

计、研发的开源分布式关系型数据库 TiDB，分布式事务型的键值数据库 TiKV，文档型项目 docs-cn、docs 等。其次是阿里，TOP 10 项目中上榜的项目有 2 个，分别是蚂蚁金服采用 React 封装的一套组件库 Ant Design（位于第 2 名），以及致力于配置和管理微服务的特性集 Nacos。百度也上榜了 2 个项目，主要集中在人工智能领域，是百度深度学习平台 PaddlePaddle 的 2 个项目，分别是核心框架 Paddle 以及相关工具库。

图表 31：GitHub2020 中国项目活跃度 TOP20

| 排名 | 项目名                               | 活跃度   | 参与开发者数量 | issue comment | open issue | open pull | pull review comment | merge pull | star   | fork  |
|----|-----------------------------------|-------|---------|---------------|------------|-----------|---------------------|------------|--------|-------|
| 1  | pingcap/tidb                      | 210.1 | 5,831   | 53,022        | 2,801      | 4,969     | 10,928              | 3,459.2    | 4,862  | 1,052 |
| 2  | ant-design/ant-design             | 193.3 | 23,620  | 32,026        | 4,836      | 3,131     | 3,320               | 2,130.7    | 12,709 | 8,052 |
| 3  | PaddlePaddle/Paddle               | 127.4 | 4,842   | 15,329        | 2,256      | 5,656     | 9,625               | 3,478.2    | 3,574  | 786   |
| 4  | tikv/tikv                         | 81.7  | 2,593   | 17,817        | 997        | 2,019     | 5,547               | 1,279.9    | 2,129  | 434   |
| 5  | apache/shardingsphere             | 75.3  | 5,267   | 9,055         | 1,713      | 3,235     | 1,858               | 2,539.5    | 3,834  | 1,443 |
| 6  | apache/incubator-tvm              | 70.4  | 2,148   | 7,961         | 437        | 2,112     | 8,506               | 1,540.1    | 1,454  | 662   |
| 7  | pingcap/docs-cn                   | 65.1  | 532     | 8,202         | 96         | 2,965     | 6,959               | 2,315.9    | 140    | 320   |
| 8  | apache/incubator-echarts          | 64.2  | 11,638  | 7,650         | 1,620      | 324       | 346                 | 194.5      | 6,664  | 4,463 |
| 9  | pingcap/pd                        | 60.9  | 437     | 13,325        | 667        | 1,667     | 4,972               | 1,297.7    | 214    | 224   |
| 10 | alibaba/nacos                     | 59.9  | 9,956   | 7,042         | 1,640      | 706       | 827                 | 410.0      | 6,347  | 3,450 |
| 11 | NervJS/taro                       | 54.7  | 7,469   | 9,339         | 2,231      | 917       | 135                 | 551.5      | 5,250  | 1,012 |
| 12 | youzan/vant                       | 54.2  | 9,806   | 4,897         | 1,661      | 715       | 201                 | 554.4      | 4,672  | 4,502 |
| 13 | pingcap/docs                      | 53.9  | 314     | 7,014         | 64         | 2,736     | 5,226               | 2,257.8    | 90     | 164   |
| 14 | ElemeFE/element                   | 52.7  | 11,749  | 4,993         | 1,762      | 297       | 10                  | 33.3       | 6,853  | 3,411 |
| 15 | apache/skywalking                 | 51.9  | 5,556   | 6,783         | 1,084      | 860       | 3,455               | 583.4      | 4,201  | 1,471 |
| 16 | PaddlePaddle/PaddleOCR            | 47.9  | 9,394   | 4,039         | 1,033      | 573       | 622                 | 420.0      | 8,430  | 1,664 |
| 17 | apache/incubator-dolphinscheduler | 47.1  | 2,588   | 9,364         | 1,269      | 1,407     | 730                 | 902.7      | 1,835  | 909   |
| 18 | apache/apisix                     | 45.4  | 2,923   | 5,855         | 1,109      | 1,029     | 3,383               | 715.0      | 2,496  | 579   |
| 19 | seata/seata                       | 45.1  | 7,339   | 3,754         | 785        | 517       | 1,805               | 313.5      | 5,261  | 2,296 |
| 20 | pingcap/tidb-operator             | 45.1  | 425     | 8,627         | 703        | 1,498     | 3,683               | 1,172.1    | 240    | 140   |

资料来源：中国开源软件推进联盟，GitHub

## 4.3 资本市场上投资人对开源领域布局力度加大

在中国由于越来越多的开源企业正在积极地进行商业化，大量资本投资方也看准了投资机会，进入开源市场。目前国内比较代表性的开源项目投融资情况表现出具有发展潜力的开源项目往往呈现出实力资本多轮注资，从早期开始长期陪伴企业发展的现象，如 PingCAP(TiDB)、巨杉数据库 (SequoiaDB) 等。表明企业在融资后利用资金进行了有效发展和扩张，融资后路径满足资本预期。2020 年和 2021 年是开源市场迎来资本热潮的时期，大部分公司在 2020 年至 2021 年初都迎来了最新融资，例如 PingCAP 在 2020 年完成 D 轮 2.7 亿美元融资创造了全球数据库历史新的里程碑，以及 Zilliz 创下了全球开源基础软件领域迄今为止最大单笔 B 轮融资的记录。

图表 32：国内开源软件企业融资状况

| 公司           | 开源核心                  | 最新一轮<br>融资金额 | 最新一轮<br>融资时间 | GitHub<br>Star | GitHub<br>Fork | Github<br>Contributor |
|--------------|-----------------------|--------------|--------------|----------------|----------------|-----------------------|
| 支流科技         | Apache APISIX         | 数百万美元        | 2021/6       | 6.5k           | 1.3k           | 228                   |
| 偶数科技         | Apache HAWQ           | 近 2 亿人民币     | 2021/8       | 619            | 304            | 69                    |
| 跬智信息技术       | Apache Kylin          | 7000 万美元     | 2021/4       | 3.2k           | 1.4k           | 184                   |
| StreamNative | Apache Pulsar         | 2300 万美元     | 2021/10      | 9.7k           | 2.4k           | 452                   |
| SphereEx     | Apache ShardingSphere | 数百万美元        | 2021/5       | 14.6k          | 5.1k           | 292                   |
| 映云科技         | EMQ                   | 1.5 亿人民币     | 2018/4       | 8.6k           | 1.6k           | 75                    |
| 泛化智能         | GAAS                  | 千万级          | 2018/10      | 1.5k           | 382            | 11                    |
| Gitee        | git                   | 未披露          | 2020/8       | 40k            | 22k            | 408                   |
| 极纳科技         | Jina                  | 3000 万美元     | 2021/11      | 11.5k          | 1.3k           | 117                   |
| Juicedata    | JuiceFS               | 数百万美元        | 2018/10      | 3.8k           | 292            | 34                    |
| 才云科技         | Kubernetes            | 数千万元         | 2019/12      | 82k            | 30k            | 373                   |
| Zilliz       | milvus                | 4300 万美元     | 2020/11      | 8.1k           | 1.1k           | 141                   |
| 欧若数网         | Nebula                | 近千万美元        | 2020/11      | 6.6k           | 668            | 61                    |
| 一流科技         | oneflow               | 5000 万元      | 2021/2       | 2.7k           | 307            | 92                    |
| EasyStack    | OpenStack             | 未披露          | 2021/1       | 4.1k           | 1.5k           | 357                   |
| 好雨科技         | Rainbond              | 未披露          | 2016/11      | 2.8k           | 565            | 18                    |
| 睿赛德科技        | RT-Thread             | 未披露          | 2020/1       | 6.3k           | 3.6k           | 309                   |
| 巨杉数据库        | SequoiaDB             | 数亿元人民币       | 2020/10      | 251            | 103            | 0                     |
| 涛思科技         | TDengine              | 4700 万美元     | 2021/5       | 16k            | 4k             | 123                   |
| 顶想信息         | ThinkPHP              | 100 万人民币     | 2015/3       | 7.7k           | 1.7k           | 37                    |
| PingCAP      | TiDB                  | 未披露          | 2021/7       | 29k            | 4.7k           | 398                   |
| 数字天堂         | uni-app               | 未披露          | 2018/9       | 34k            | 3.1k           | 170                   |
| 云轴信息         | ZStack                | 未披露          | 2021/3       | 1.1k           | 364            | 43                    |
| 易软天创         | 禅道                    | 未披露          | 2021/10      | 777            | 239            | 35                    |
| Chatopera    | 春松客服                  | 数百万美元        | 2018/8       | 1,9k           | 623            | 3                     |
| 泽拓科技         | 昆仑数据库                 | 数千万元人民币      | 2021/8       | 87             | 13             | 6                     |
| 云联壹云         | 云联壹云                  | 未披露          | 2021/1       | 661            | 187            | 18                    |

资料来源：公司公告，Crunchbase，云启资本；统计时间截至 2021 年 12 月 31 日

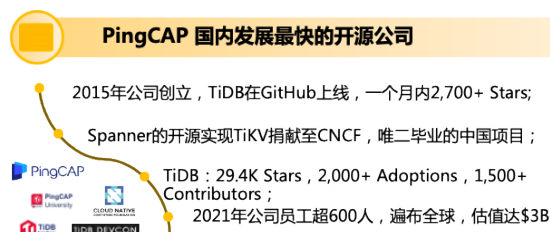
## 5 案例分析

### 5.1 PingCAP

PingCAP 国内领先的企业级开源分布式数据库厂商。公司成立于 2015 年，致力于为全球行业用户提供

稳定高效、安全可靠、开放兼容的新型数据基础设施。核心产品为分布式关系型数据库 TiDB。TiDB 是 PingCAP 公司自主设计、研发的开源分布式关系型数据库，是一款同时支持在线事务处理与在线分析处理 (Hybrid Transactional and Analytical Processing, HTAP) 的融合型分布式数据库产品。

图表 33 : PingCAP 发展历史



资料来源：公司公告，GitHub，云启资本

图表 34 : PingCAP 核心产品 TiDB



资料来源：公司公告，云启资本

PingCAP 的发展潜力受到资本追捧。PingCAP 在 2020 年底完成了 2.7 亿美元的 D 轮融资，成功迈入独角兽行列也成为中国开源新势力的领军人物。2021 年 7 月，公司完成完成估值 30 亿美元的新一轮融资。在 PingCAP 的融资历程中，出现了资本连续加持的情况。经纬中国从 2015 年的天使轮投起，连投了 5 轮；云启资本从 2016 年的 A 轮进入，也跟进到最新一轮。

图表 35: PingCAP 融资历程

| 日期         | 融资轮次 | 融资金额     | 投资机构                                                                                              |
|------------|------|----------|---------------------------------------------------------------------------------------------------|
| 2021-07-20 | 战略融资 | 未披露      | 红杉基金，GIC，贝塔斯曼亚洲投资基金，纪源资本，五源资本                                                                     |
| 2020-11-17 | D 轮  | 2.7 亿美元  | Coatue，云启资本，FutureX 天际资本，贝塔斯曼亚洲投资基金，时代资本，经纬中国，挚信资本，纪源资本，昆仑资本，五源资本，Access Technology Ventures，晨曦投资 |
| 2018-09-12 | C 轮  | 5000 万美元 | 云启资本，复星创富，经纬中国，华创资本，五源资本                                                                          |
| 2017-06-13 | B 轮  | 1500 万美元 | 峰瑞资本，云启资本，经纬中国，险峰长青，华创资本                                                                          |
| 2016-09-21 | A 轮  | 700 万美元  | 峰瑞资本，云启资本，初心资本，经纬中国                                                                               |
| 2015-12-09 | 天使轮  | 数百万元人民币  | 经纬中国                                                                                              |

资料来源：公司公告，Crunchbase，云启资本

PingCAP 产品 - 社区 - 商业化生态处于全国领先。PingCAP 产品具备一键水平扩容或者缩容、金融级高可用、实时 HTAP、云原生的分布式数据库、高度兼容 MySQL 生态等重要特性。截止到 2021 年 11 月，TiDB 项目在 GitHub 上已累计获得超过 29,000 颗星，近 1,500 位开源代码贡献者，达到全球



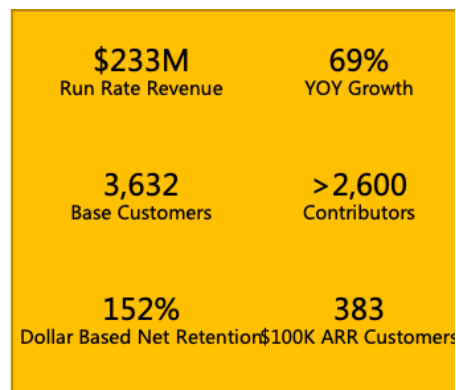
年，总部位于特拉华州。最初起源于一个个人项目，是由 GitLab 前 CTO Dmitriy 2001 年创建的。根据 GitLab 2021 年的上市招股说明书所示，GitLab 截至 2021 年 7 月 31 日，公司年化营收超过 2.3 亿，最近的 QRR 达到 5,800 万美元，同比增长 69%。GitLab 拥有 ARR 大于 5,000 美元的客户 3,632 个，ARR 大于 10 万美元的客户 383 个，包括高盛、西门子、英伟达等，开源社区中的代码贡献者超过 2,600 位。

图表 37 : GitLab 发展历史



资料来源：公司公告，云启资本

图表 38 : GitLab 经营状况



资料来源：Pitchbook，云启资本

GitLab 采用的是 OpenCore 开放核心的开源商业模式。服务共分为三级，包括一个免费套餐与两个收费套餐，分别针对社区开源用户、中小企业与大型企业。免费套餐主推面向大众用户的 Create 和 Verify 模块，两个收费套餐分别增加了版本控制、运营分析、项目管理与进阶安全测试、产品组合管理、合规规划等内容。订阅模式分为建立在公有云上的 SaaS 服务与企业本地管理，在企业本地管理的情况下，企业可以根据情况将 GitLab DevOps 平台建立在私有云或者混合云上，同时 GitLab 也会将本地管理的控制权完全交予企业。

图表 39 : GitLab 的 Open Core 商业模式

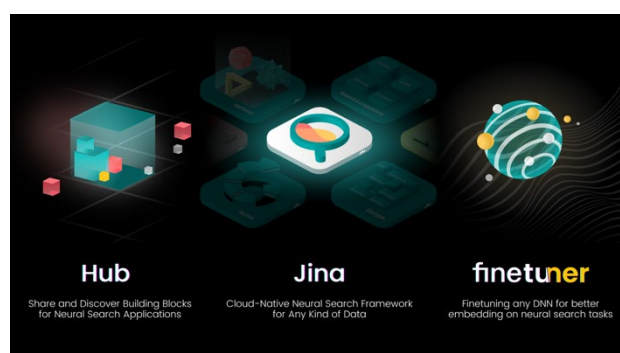
| Open Core模式，三条产品线          |                                                                       |
|----------------------------|-----------------------------------------------------------------------|
| 标准版GitLab<br>(针对个人用户)      | DevOps全生命周期所有阶段；自带GitLab CI运行器；支持任意生产环境部署；免费静态网站；获取每月400分钟的CI/CD流水线时间 |
| 专业版GitLab<br>(提高团队协作)      | 更好的代码评审；运维可视化；项目管理 发布控制；SaaS用户每月10,000分钟CI/CD流水线时间                    |
| 旗舰版GitLab<br>(企业级安全&合规&规划) | 高级安全测试；合规性；项目组合管理；价值流分析 访客不计入计费用户；SaaS用户每月50,000分钟CI/CD流水线时间          |

资料来源：公司公告，云启资本

## 5.3 Jina AI

Jina 首创的“神经搜索”，可以让企业利用可操作的非结构化数据构建搜索解决方案，做出更有效的业务决策。Jina AI 的核心项目在 Github 上被称为 Jina，允许用户在短短几分钟内创建一个由深度学习驱动的云原生搜索解决方案。Jina 将构建一个生产就绪的神经搜索系统的工作量从几个月减少到几分钟，非常适合需要快速、轻量级开发周期的商业环境。除了主产品 Jina, 公司最近推出了 Finetuner, 它可以让用户根据企业的独特需求对神经搜索系统进行调整。

图表 40 : Jina 产品矩阵

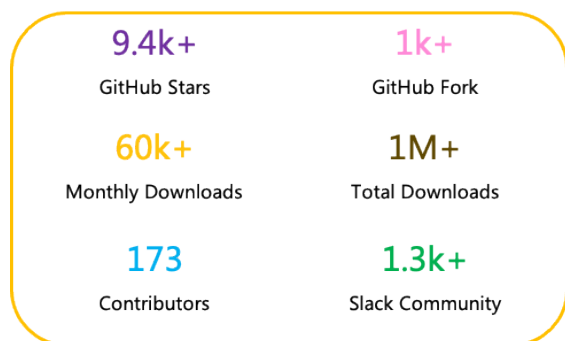


资料来源：Jina 官网

Jina AI 已经受到社区的积极认可，并且受到资本不断注入。自 2020 年 2 月成立至今，公司已经创建了超过 1,000 个用户的大规模开发者社区。Jina 仅用了 18 个月的时间在 GitHub 上吸引了大量用户，GitHub stars 达到 9.4k+, fork 数达到 1k+, 总下载量超过 100 万次。Jina 用户场景多样，可以满足不同行业对神经搜索的需求，比如游戏制作公司的 3D 模型，电商网站上的图像和能够理解混合语义的

问答聊天机器人。Jina 在资本市场上也受到多个知名投资机构认可，目前公司总融资金额已达到 3,900 万美元，最新 A 轮融资 3,000 万美元，该轮融资由新投资方美国投资机构 Canaan Partners 领投，Mango Capital 以及原投资方 GGV 纪源资本、SAP.iO 和云启资本跟投。

图表 41：Jina GitHub 运营情况



资料来源：GitHub，云启资本

图表 42：Jina AI 融资历史

| 披露日期       | 交易金额     | 融资轮次    | 投资方                                                          |
|------------|----------|---------|--------------------------------------------------------------|
| 2021-11-22 | 3000 万美元 | A 轮     | Canaan Partners, Mango Capital, GGV 纪源资本, SAP, iO Fund, 云启资本 |
| 2020-09-22 | 600 万美元  | Pre-A 轮 | GGV 纪源资本, SAP, iO Fund, 云启资本                                 |
| 2020-05-13 | 500 万美元  | 种子轮     | 云启资本                                                         |

资料来源：Crunchbase，云启资本

## 5.4 Zilliz

Zilliz 是研发面向人工智能的新一代数据处理和分析平台，其主要是为应用型企业提供底层技术。Zilliz 研发的 GPU 加速的 AI 数据中台解决方案 Mega，其中包括数据 ETL 系统 MegaETL、数据库系统 MegaWise、面向 Hadoop 生态的模型训练系 MegaLearning 和特征向量检索系统 Milvus，可满足传统的加速数据 ETL、加速数据仓库和加速数据分析的场景和需求，面向各类新兴的 AI 应用场景，已被全球 1,000 多个企业使用，涵盖金融、电信、安防、智慧城市和电子商务等行业。

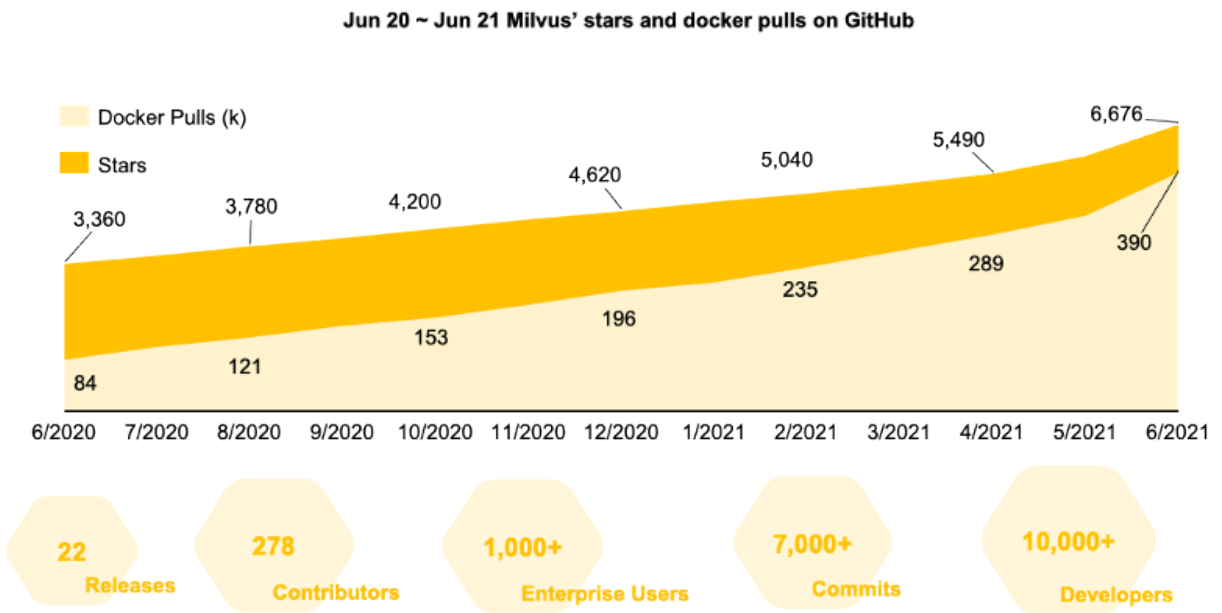
图表 43：Zilliz 全球用户

| 1,000+ Global enterprise users |                    |                |                                |                             |               |                   |  |
|--------------------------------|--------------------|----------------|--------------------------------|-----------------------------|---------------|-------------------|--|
| COMPASS                        | AXA                | OPERA software | dailyhunt                      | MessageBird                 | moj           | SmartNews         |  |
| Lucidworks                     | FINNOS             | identiv        | tokopedia                      | LINE                        | TREND MICRO   | WISEO BITS        |  |
| PrivatBank                     | Neoway             | acesso digital | THE CLEVELAND MUSEUM OF ART    | Haier                       | Hisense       | TRANSSION 珍奇控股    |  |
| 快手                             | mi                 | 爱奇艺            | 蜻蜓FM                           | 喜马拉雅                        | 中国电信          | 科大讯飞              |  |
| 企查查 qichacha.com               | Alibaba Group      | 永辉超市           | 丁香园 WWW.DXY.CN                 | 泰康人寿 Talking Life           | 阳光 阳光出行       | 什么值得买 让消费更简单      |  |
| 唯品会 品牌特卖                       | 中国太平 CHINA TAIPING | 用友 yonyou      | 中国建设银行 China Construction Bank | 交通银行 BANK OF COMMUNICATIONS | 爱建信托 AJ TRUST | 鹏华基金 PENGHUA FUND |  |
| 平安银行 PING AN BANK              | 招商银行 信用卡           | 峰网电力 OR POWER  | 新浪微博 weibo.com                 | W MPS                       | 汽车之家 看车·买车·用车 | 紫光华智 UNISINSIGHT  |  |

资料来源：公司公告，云启资本

Zilliz 的成功代表着基于 GPU 的大数据加速器为企业日益增长的数据分析需求提供了有效解决方案。Zilliz 的核心项目向量相似度搜索引擎 Milvus 是全球首款 GPU 加速海量特征向量匹配和检索引擎。Milvus 依托 GPU 加速，提供极速特征向量匹配以及多维度数据联合查询（特征、标签、图片、视频、文本和语音等联合查询）功能，并且支持自动分库分表和多副本，能对接 TensorFlow、PyTorch 和 MxNet 等 AI 模型，可实现百亿特征向量的秒级查询。Milvus 于 2019 年 10 月在 GitHub 上开源，Stars 和 Docker Pulls 数量持续高速增长，2021 年 6 月达到 6,000+，拥有近 300 位贡献者和 10,000+ 用户的开发者社区。资本市场上，Zilliz 在 B 轮获 4,300 万美金，成为全球开源基础软件最大单笔 B 轮融资，表明了投资机构对 Zilliz 未来发展潜力的看好。

图表 44：Zilliz GitHub 社区运营情况



资料来源：GitHub，云启资本

## 5.5 EMQ 映云科技

EMQ 公司是一家开源物联网数据基础设施软件供应商。EMQ 开源项目发起于 2013 年，2017 年在杭州正式成立公司，致力于提供覆盖云边端的开源物联网消息中间件及流数据库产品，一站式满足企业在 5G 时代物联网实时数据移动分发、流处理及分析的需求。有超过 10 个开发团队，分布在全球各地。目前，EMQ 国内在北京、上海、深圳、南京、昆明设有分支机构；海外研发总部设在斯德哥尔摩，在瑞典、德国、北美和日本设有分支机构或服务团队。

图表 45 : EMQ 公司情况概览



资料来源：EMQ 官网

EMQ 是全球开源 MQTT 消息服务器领导者。EMQ 目前在全球拥有近 1 万家开源用户、近 400 家企业客户和 20 多家世界五百强合作伙伴。公司商业客户与合作伙伴来自多个国家和地区，包括 HPE、Vmware、Ericsson、Verifone、Telstra、Nokia、华为、中国电信、中国移动、上汽大众、中国银联、国家电网、台积电等海内外大型企业，从车联网到工业互联网，从运营商到金融支付，从能源电力到智慧城市构建了一站式物联网平台与应用。

图表 46 : EMQ 业务情况



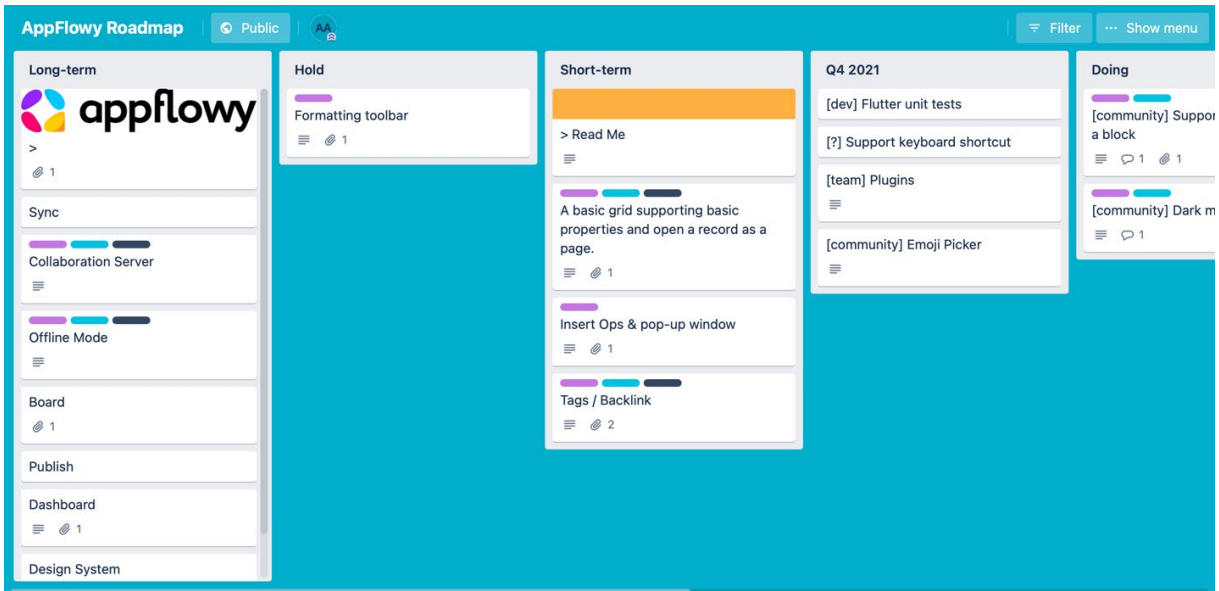
资料来源：EMQ 官网

## 5.6 AppFlowy

针对 Notion 数据安全、移动端适配等局限性，AppFlowy 意在成为 Notion 的开源替代。AppFlowy 的开源模式使其具备一定的优势：首先，用户可以实现 100% 数据控制。尤其是对于团队用户来说，可以

随时随地托管 AppFlowy，没有供应商锁定的问题。其次，向用户提供了定制与扩展功能。企业或者团队用户可以使用开放的核心代码库以定制的方式设计和修改 AppFlowy；无编码经验的个人用户可以使用社区驱动的工具箱，包括模板、插件、主题等。再次，由于 AppFlowy 是基于 Flutter 和 Rust 构建的，因此可以很好的支持多平台和设备上的使用。

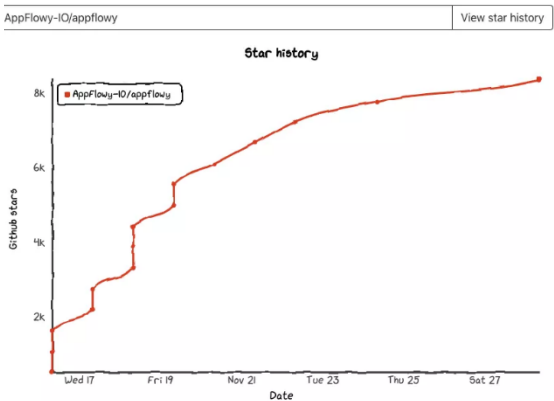
图表 47 : AppFlowy Roadmap



资料来源：AppFlowy 官网

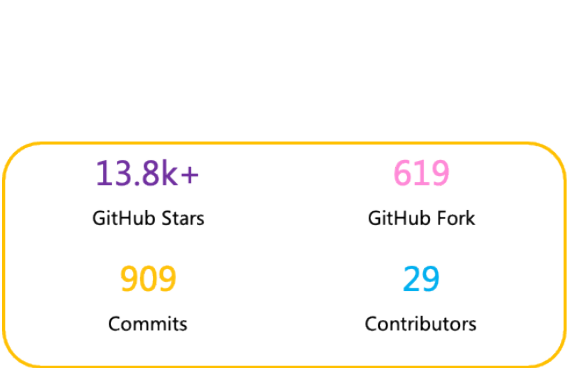
AppFlowy 一经发布就受到广泛的关注。在 GitHub 上线的短短一周就获得了近 8k Star。而截至 2021 年 12 月 25 号的最新数据，AppFlowy 的 GitHub stars 达到了 13.8k+，Fork 数达到 619，贡献者人数达到 29 人。

图表 48 : AppFlowy 上线一周内 Github 运营情况



资料来源：AppFlowy, GitHub

图表 49 : AppFlowy 最新 Github 运营情况



资料来源：GitHub, 云启资本

## 5.7 Confluent

2021 年 6 月 Confluent 在纳斯达克 IPO 上市，估值高达 114 亿美元。Confluent 的核心产品是 Apache Kafka。Kafka 是由 Jay Kreps、Jun Rao 和 Neha Narkhede 于 2011 年在 LinkedIn 内部创建，2014 年，三位创始人将公司独立拆分出来。Confluent 的 Apache Kafka 是一个高吞吐量的分布式发布订阅消息系统，用于在技术系统之间传输信息，已经拥有超过 6 万名来自全球的社区成员，大约超过 70% 的财富 500 强企业使用过 Kafka，包括花旗集团、Humana、英特尔和沃尔玛等。

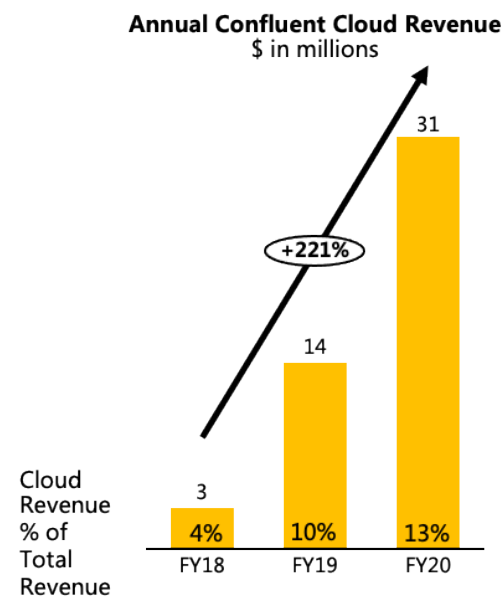
图表 50 : Confluent 不同行业的企业用户



资料来源：Confluent 官网

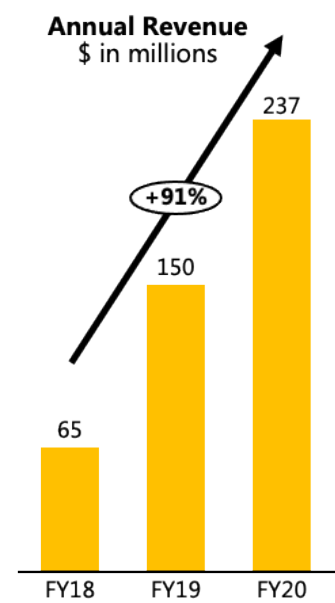
Confluent 提供的云服务是营业收益的一项主要来源。过去三年，Confluent 的销售额年均增长率为 90%，而其中云服务的收入占比从 2018 财年的 4% 升至 2020 财年的 13%，复合增长率高达 221%，超过其总收入的增长率，可以看出 Confluent 很看重云市场，下一步将重点发展 Confluent Cloud。Confluent 通过三个公共云提供 Confluent Cloud：Microsoft、AWS 和 Google。

图表 51 : Confluent FY18 至 FY20 云服务年收益



资料来源：Pitchbook，云启资本

图表 52 : Confluent FY18 至 FY20 总年收入



资料来源：Pitchbook，云启资本

## 5.8 MongoDB

MongoDB 是全球领先的 NoSQL 数据库平台。当前产品主要包括企业级数据库 产品 MongoDB Enterprise Advanced、云数据库产品 MongoDB Atlas、开源数据库产品 Community Server 和移动数据库及同步平台 MongoDB Realm。公司对于不同产品采用不同的定价方式，企业版按服务器节点订阅，云产品按需付费。以云数据库产品 MongoDB Atlas 为例，公司面向小型团队提供共享集群，共享内存和算力，并根据不同的存储空间按使用时长计费；面向专业开发团队提供专属集群，根据不同的 RAM、算力和存储空间按使用时长计费。

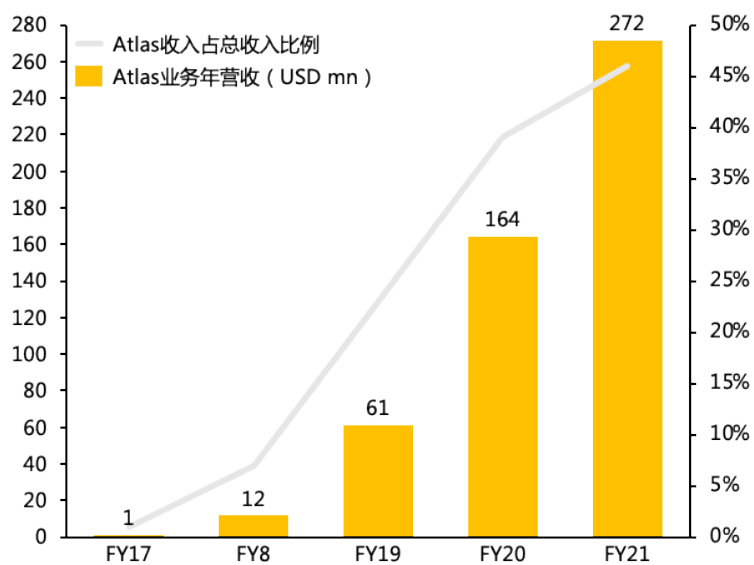
图表 53 : Atlas Dedicated Cluster 版本定价规则

| Cluster | Storage | RAM    | vCPUs    | Base Price |
|---------|---------|--------|----------|------------|
| M10     | 10 GB   | 2 GB   | 2 vCPUs  | \$0.08/hr  |
| M20     | 20 GB   | 4 GB   | 2 vCPUs  | \$0.20/hr  |
| M30     | 40 GB   | 8 GB   | 2 vCPUs  | \$0.54/hr  |
| M40*    | 80 GB   | 16 GB  | 4 vCPUs  | \$1.04/hr  |
| M50*    | 160 GB  | 32 GB  | 8 vCPUs  | \$2.00/hr  |
| M60*    | 320 GB  | 64 GB  | 16 vCPUs | \$3.95/hr  |
| M80*    | 750 GB  | 128 GB | 32 vCPUs | \$7.30/hr  |
| M140    | 1000 GB | 192 GB | 48 vCPUs | \$10.99/hr |
| M200*   | 1500 GB | 256 GB | 64 vCPUs | \$14.59/hr |
| M300*   | 2000 GB | 384 GB | 96 vCPUs | \$21.85/hr |
| M400*   | 3000 GB | 488 GB | 64 vCPUs | \$22.40/hr |
| M700    | 4000 GB | 768 GB | 96 vCPUs | \$33.26/hr |

资料来源: MongoDB 官网

MongoDB 云业务快速发展。Atlas 是 MongoDB 托管的多云产品, MongoDB Atlas Free Tier 产品现已在北美、欧洲和亚太地区的所有三大云提供商(亚马逊网络服务(“AWS”)、谷歌云平台(“GCP”)和微软(Azure))上提供, 受益于全球数据量(尤其是非结构数据)的爆发式增长、本地数据库向云端的不断迁移, MongoDB 营收持续高速增长。自 2017 财年 MongoDB Atlas 发布以来, Atlas 贡献的收入占比由当年的 1% 迅速上升至 2021 财年前三季度的 44%。2021 财年, Atlas 产品实现收入 2.72 亿美元, 同比增加 46%。

图表 54 : MongoDB Atlas 业务情况



资料来源: Pitchbook, 云启资本

## 附录 1 开源硬件 v.s. 开源软件

开源硬件，即 Open Source Hardware，是可以通过公开渠道获得的硬件设计，如电路图、材料清单和开发板布局数据。并且硬件设计的源代码的特定的格式可以为其他人获得，以方便对其进行修改。以 CPU 处理器为例，开源芯片的开源通常包括三个层次，首先是指令集（ISA）开源，例如 RISC-V 指令集规范是开源的；其次是微架构的设计开源；第三层是开源 RTL 源码乃至 GDSII，例如 Berkeley 的 Rocket Chip、剑桥大学的 lowRISC、芯来蜂鸟 E203 等都开源了 RTL 级源码。

### 专家点评

段夕华：开源硬件尤其是 RISC-V 前途不可限量

就开源硬件而言，与开源软件概念真正可比的是第三层的硬件开源，即开放 RTL 级源码。但硬件代码和软件代码有较大的差异，从硬件 RTL 代码到 GDSII 仍有较长的开发流程，而没有经过流片验证的 GDSII 也并不具备实际的参考价值。因此考虑到软件和硬件开发周期的不同、迭代成本的差异，硬件 RTL 级源码开源后，能撬动的开发杠杆较为有限，目前来看市面上只有少数学术或社区推广性质的产品。2021 年 12 月首届滴水湖中国 RISC-V 产业论坛上，与会的专业嘉宾投票结果显示，超 60% 的嘉宾认为 RISC-V 芯片还不应该被开源。

## 附录 2 美国开源资本市场情况

| COSS Company              | FOSS Core     | VC Raised(M) | VC Date | GitHub Star | GitHub Fork | Github Contributor |
|---------------------------|---------------|--------------|---------|-------------|-------------|--------------------|
| Alfresco                  | Alfresco      | \$70         | 2014/8  |             |             |                    |
| Segment                   | Analytics.js  | \$284        | 2020/1  | 4.7k        | 766         | 59                 |
| Jfrog                     | Artifactory   | \$227        | 2018/10 | 828         | 1.9k        | 57                 |
| Instructure               | Canvas        | \$90         | 2015/2  | 4.1k        | 1.7k        | 248                |
| Datastax                  | Cassandra     | \$228        | 2021/5  | 1.8k        | 868         | 95                 |
| Pivotal(Now VMware Tanzu) | CloudFoundry  | \$1,700      | 2016/5  | 2k          | 648         | 303                |
| Couchbase                 | Couchbase     | \$251        | 2020/5  | 1.5k        | 291         | 48                 |
| Acquia                    | Drupal        | \$174        | 2018/8  |             |             |                    |
| Elastic                   | ElasticSearch | \$162        | 2018/5  | 56.7k       | 20.7k       | 1668               |
| Mozilla Corporation       | Firefox       | \$22         |         | 10.5k       | 2.4k        | 209                |
| Treasure Data             | Fluentd       | \$54         | 2016/11 | 10.6k       | 1.2k        | 220                |
| GitLab                    | Git           | \$415        | 2020/11 | 40k         | 22k         | 408                |
| GitHub                    | Git           | \$350        | 2015/12 | 39.7k       | 22.3k       | 1493               |

|                    |                         |         |         |       |       |       |
|--------------------|-------------------------|---------|---------|-------|-------|-------|
| Cloudera           | Hadoop                  | \$1,040 | 2018/4  | 12k   | 7.4k  | 395   |
| JetBrains          | IntelliJ                | –       |         | 12.7k | 4.2k  | 717   |
| Cloudbees          | Jenkins                 | \$111   | 2018/6  | 17.9k | 7k    | 673   |
| Confluent          | Kafka                   | \$455   | 2020/4  | 2.5k  | 675   | 85    |
| Kaltura            | Kaltura                 | \$166   | 2016/8  | 1.5k  | 349   | 21    |
| Liferay            | Liferay Portal          | –       |         | 1.8k  | 3k    | 700   |
| VA Linux(Geeknet)  | Linux                   | \$30    |         |       |       |       |
| Red Hat            | Linux                   | \$5     |         | 120k  | 39.3k | 5000+ |
| SUSE               | Linux Kernel            | –       |         | 16    | 5     | 5000+ |
| Magento Commerce   | Magento                 | \$273   | 2017/1  | 9.7k  | 8.6k  | 1475  |
| HashiCorp          | Many                    | \$349   | 2020/3  |       |       |       |
| Mapbox             | Mapbox GL JS            | \$334   | 2020/5  | 7.9k  | 1.8k  | 311   |
| Rapid7             | Metasploit              | \$89    | 2014/12 | 25.3k | 11.6k | 901   |
| Docker             | Moby                    | \$330   | 2021/5  | 61.3k | 17.7k | 2132  |
| MongoDB(fka 10gen) | MongoDB                 | \$311   | 2018/3  | 20.5k | 5k    | 541   |
| MuleSoft           | Mule ESB                | \$259   | 2016/3  | 245   | 625   | 116   |
| MySQL AB           | MySQL                   | \$40    | 2006/2  | 7k    | 2.7k  | 86    |
| Neo4j              | Neo4j                   | \$516   | 2021/6  | 9.4k  | 2k    | 213   |
| Odoo               | Odoo                    | \$319   | 2021/7  | 22.9k | 15.1k | 1275  |
| Nicira             | Open vSwitch            | \$42    | 2011/2  | 2.7k  | 1.6k  | 410   |
| ForgeRock          | OpenAM/IDM/DJ/IG        | \$234   | 2020/4  | 105   | 54    | 15    |
| Mirantis           | OpenStack               | \$227   | 2015/8  | 4.1k  | 1.5k  | 357   |
| Rackspace          | OpenStack               | \$227   | 2017/9  | 4.1k  | 1.5k  | 357   |
| Pentaho            | Pentaho                 | \$75    | 2012/10 | 5.3k  | 2.7k  | 185   |
| Enterprise DB      | Postgres                | \$68    | 2011/2  | 9k    | 3k    | 48    |
| Postman            | Postman Runtime         | \$433   | 2021/8  | 133   | 75    | 32    |
| Puppet Labs        | Puppet                  | \$190   | 2020/7  | 6.3k  | 2.2k  | 561   |
| Redis Labs         | Redis                   | \$356   | 2021/4  | 51.4k | 20.1k | 524   |
| Linden Lab         | Second Life Viewer      | \$19    | 2006/3  |       |       |       |
| Sourcefire         | Snort                   | \$40    | 2006/10 | 1.2k  | 353   | 18    |
| Databricks         | Spark                   | \$3,500 | 2021/8  | 31.1k | 24.6k | 1728  |
| SugarCRM           | SugarCRM                | \$123   | 2018/8  |       |       |       |
| Talend             | Talend Data Integration | \$102   | 2013/12 |       |       |       |
| Canonical          | Ubuntu                  | \$13    | 2013/8  |       |       |       |
| Fastly             | Varnish                 | \$220   | 2018/7  | 2.7k  | 336   | 79    |
| WP Engine          | Wordpress               | \$291   | 2018/1  |       |       |       |
| Automattic         | Wordpress               | \$986   | 2021/2  |       |       |       |

## 附录 3 YC 开源项目一览

| COSS Company          | FOSS Core    | Space                       | YC Batch | Funding (M) |
|-----------------------|--------------|-----------------------------|----------|-------------|
| Activeloop            | Activeloop   | Data Pipelines              | S18      | \$2         |
| Data Mechanics        | Apache Spark | Data Science                | W19      | –           |
| BackType              | Apache Storm | Data Analytics              | S08      | \$1         |
| Athens                | Athens       | Note-Taking App             | W21      | –           |
| 280 North             | Cappuccino   | Web Application Development | W08      | –           |
| Chatwoot              | Chatwoot     | Helpdesk / "Modern ITSM"    | W21      | –           |
| Dataform              | Dataform     | Data Engineering            | W18      | –           |
| DevicePlane           | Deviceplane  | Embedded Systems            | W20      | –           |
| Docker / fka DotCloud | Docker       | DevOps                      | S10      | \$308       |
| AppJet                | Etherpad     | Document Editing            | S07      | \$1         |
| kSense (Jitsu)        | EventNative  | Data Collection / Analytics | S20      | –           |
| Fig                   | Fig          | DevTools                    | S20      | –           |
| Flynn                 | Flynn        | DevOps                      | S14      | –           |
| GitDuck               | GitDuck      | DevTools                    | S20      | –           |
| GitLab                | GitLab/Git   | DevOps                      | W15      | \$436       |
| Manycore              | GraalVM      | Developer Tool              | S20      |             |
| Apollo fka Meteor     | GraphQL      | DevOps                      | S11      | \$53        |
| Influx Data           | InfluxDB     | Database                    | W13      | \$120       |
| Insoshi               | Insoshi      | Social Network              | W08      | –           |
| Protocol Labs         | IPFS         | Storage                     | S14      | \$260       |
| EQ Alpha Tech         | KeyDB        | Database                    | S20      |             |
| CoreOS                | Kubernetes   | DevOps                      | S13      | \$48        |
| KubeSail              | Kubernetes   | DevOps                      | S19      | –           |
| Lunatic               | Lunatic      | W21                         | –        |             |
| Mattermost            | Mattermost   | Messaging Application       | S12      | \$70        |
| MindsDB               | MindsDB      | Machine Learning            | W20      | –           |
| FathomDB              | MySQL        | Database Hosting            | W08      | –           |
| Heroic Labs           | Nakama       | Game Engine                 | S15      | –           |
| Flotype (Bridge)      | NowJS        | DevOps                      | W11      | \$2         |
| Okteto                | Okteto       | DevOps                      | W19      | –           |
| OneGraph              | OneGraph     | Developer Platform          | W18      | –           |
| Ukama                 | OpenCellular | Cellular Networks           | S20      | –           |
| Clickpass             | OpenID       | Authentication / Identity   | S07      | –           |
| Opentrons             | Opentrons    | Robotics                    | W16      | \$40        |
| Optic                 | Optic        | API Platform                | W18      | –           |

|                             |                     |                              |     |      |
|-----------------------------|---------------------|------------------------------|-----|------|
| Our World In Data           | OWID                | Data Portal                  | W19 | –    |
| Pachyderm                   | Pachyderm           | Data Science                 | W15 | \$12 |
| Papercups                   | Papercups           | Customer Service Application | S20 | –    |
| PipelineDB                  | PipelineDB          | Database                     | W14 | –    |
| Parakey                     | Planned but no OSS  | Operating System             | W07 | –    |
| Batch.sh                    | Plumber             | Middleware Messaging         | S20 | –    |
| CitusData                   | Postgres/CDB        | Database                     | S11 | \$13 |
| PostHog                     | PostHog             | Analytics                    | W20 | –    |
| Pyroscope                   | Pyroscope           | Continuous Profiling         | W21 | –    |
| QuestDB                     | QuestDB             | Database                     | S20 | –    |
| Quirk – now RoomService.dev | Quirk               | Consumer Application         | W19 | –    |
| Convex                      | Rack                | DevOps                       | S11 | –    |
| Realm                       | Realm               | Mobile Database              | S11 | \$12 |
| Replicate                   | Replicate           | Machine Learning             | W20 | \$40 |
| RethinkDB                   | RethinkDB           | Database                     | S09 | \$1  |
| Heroku                      | Ruby                | PaaS                         | S07 | \$13 |
| Sails                       | Sails               | Node.js Framework            | W15 | –    |
| SigNoz                      | SigNoz              | COSS Datadog                 | W21 | –    |
| Armory                      | Spinnaker           | DevOps                       | W17 | \$42 |
| Supabase                    | Supabase            | Database                     | S20 | –    |
| SuperTokens                 | SuperTokens         | Secure Networking            | S20 | –    |
| Synth                       | Synth               | Synthetic Data               | S20 | –    |
| NestyBox                    | SysBox              | DevOps                       | S20 | –    |
| Open Motors fka OSVehicle   | TABBY EVO           | Electric Vehicle Platform    | W16 | –    |
| Dataline                    | TBA                 | Data Integration             | W20 | –    |
| Gravitational               | Teleport/Kubernetes | DevOps                       | S15 | \$29 |
| Tipe                        | Tipe                | CMS                          | W18 | –    |
| Virtualmin                  | Virtualmin          | Web Hosting                  | W07 | –    |
| vueStorefront               | vueStorefront       | eCommerce                    | W21 | –    |
| Wasmer                      | Wasm/Wasmer         | Developer Platform           | S19 | –    |
| Wasp                        | Wasp                | Wasm                         | W21 | –    |
| Webiny                      | Webiny              | App Dev                      | W21 | –    |
| Zencoder                    | Zencoder            | Video Encoding               | W10 | \$2  |

# 2021 开源大事记

# 摘要

整理：蔡芳芳

## 一、各国开源政策将对开源世界的未来产生重大影响

2021 年，多国发布政策将开源提升至国家级别的战略高度，肯定了开源模式对信息技术创新和软件产业发展的重要性，并将繁荣开源生态作为一项重要任务，国内工信部信息技术发展司发布的《“十四五”软件和信息技术服务业发展规划》就是其中典型的一例。

## 二、开源法务合规趋势：意识增强，道阻且长

随着各行各业越来越多地使用开源代码，一些开源项目已成为“大生意”，围绕开源法务和合规话题的讨论成为焦点。中国首个明确 GPL 3.0 协议法律效力的判决案例、甲骨文诉谷歌版权侵权案尘埃落定等事件表明，软件行业对于开源法务和合规的意识正在增强。

## 三、开源治理成为显学

开源软件安全事件频发，开源软件供应链治理日趋重要。Linux 基金会积极推动 OpenChain 国内外社区接轨与交流，中国信通院与诸多国内企业密切关注或加入。此外，企业如何开源、项目开源流程、开源项目度量等问题也是产业界关注的焦点。基于量化模型的方式对社区治理效果进行各项评估，成为越来越多成熟开源组织的“标配”。Linux 基金会推动成立的度量开源项目与社区健康度的开源项目 CHAOSS，值得关注。

## 四、国际基金会的左右博弈：RMS 重回自由软件基金会与 Rust 社区争议

开源软件项目需要不同主体通力协作，社区内部的冲突无疑会对软件项目的开发和维护产生负面影响。2021 年，自由软件之父 Richard M. Stallman 重返自由软件基金会引发激烈争议、Rust 社区 Moderation Team 因不满核心团队而集体辞职等风波，都暴露出了自由和开源软件社区面临的复杂态势和治理难题。

## 五、中国开源走向世界，塑造新时代影响力

中国开发者在开源世界的影响力正在不断提升。据最新的 GitHub 年度开发者报告，2021 年中国在 GitHub 上的开发者新增了近 103 万，累计约 755 万。在 ASF、LF、CNCf 等国际开源基金会中，源自中国的开源项目越来越多，同时，开始有更多中国开源人当选国际基金会重要职位。

## 六、开源新创投资持续发光发热

2020 年底到 2021 年底，国内外基于开源项目的初创企业空前活跃，基于开源项目的商业公司获得融资、上市等现象屡见不鲜，且融资金额和估值 / 市值不断刷新上限，开源社区和开源软件的商业价值得到资本认可。

## 七、开源操作系统迎来新一轮繁荣期

2020 年底，RedHat 宣布将在 2021 年底结束对 CentOS 8 的支持，全球用户开始尝试寻找合适的操作系统进行替换，而及时解决用户在 CentOS 退出后可能面临的风险成为了操作系统厂商和研发人员努力的方向，这使得开源操作系统在 2021 年迎来新的蓬勃发展期。

## 八、Rust 迈上新征程

2021 年，Rust 基金会由 Mozilla、Amazon、华为、谷歌、微软联合成立，Rust 终于告别“动荡”。此外，Linux 社区也对 Rust 表现出了积极的接纳态度，2021 年下半年 Rust for Linux 项目进展顺利，开发者有望在 2022 年看到 Linux 内核正式支持 Rust。

## 九、AI & 低代码将会如何改变开源，值得关注

人工智能时代的开源工作，正面临着全新的挑战。今年发布的 GitHub Copilot 工具，使用了机器学习技术来提供代码建议 / 自动补全，并因此引发了不小的争议。许多开发者认为，GitHub Copilot 宣称的基于公开代码训练其实是在未遵循开源许可证的情况下，肆意“抄袭”开源代码。

## 十、开源硬件持续升温，RISC-V 成果涌现

随着市场对定制硬件需求的增多，以及越来越多初创企业开始寻求用于构建高度合适 AI/ML 算法的加速器和解决方案，开源硬件的热度持续升温。而随着 RISC-V 处理器 ISA 的出现，开源硬件已成为现实。同时，RISC-V 在中国得到了越来越多的关注和投入，也有越来越多出色成果涌现出来。

## 全文

审校（蔡芳芳、庄表伟）

作者群（按照姓氏汉语拼音排序，不分先后）：阿法兔、蔡芳芳、高原、耿航、郭雪、江炜婕、李圳虎、梁尧、刘天栋、王蓉、卫剑钊、薛亮、杨丽蕴、张俊霞、庄表伟

## 一、各国开源政策将对开源世界的未来产生重大影响

### 国内

- 2021 年 11 月 30 日，工信部信息技术发展司发布了[《“十四五”软件和信息技术服务业发展规划》](#)，规划是按照《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》编写。
- 其中提到了，开放、平等、协作、共享的开源模式，加速软件迭代升级，促进产用协同创新，推动产业生态完善，成为全球软件技术和产业创新的主导模式。当前，开源已覆盖软件开发的全域场景，正在构建新的软件技术创新体系，引领新一代信息技术创新发展，全球 97% 的软件开发者和 99% 的企业使用开源软件，基础软件、工业软件、新兴平台软件大多基于开源，开源软件已经成为软件产业创新源泉和“标准件库”。
- 在发展目标上，通知指出，生态培育获得新发展。培育一批具有生态主导力和核心竞争力的骨干企业，到 2025 年，主营业务收入达百亿级企业过百家，千亿级企业超过 15 家。建设 2-3 个有国际影响力的开源社区，培育超过 10 个优质开源项目。高水平建成 20 家中国软件名园。软件市场化定价机制进一步完善。建成一批国家特色化示范性软件学院。国际交流合作全面深化。
- 在主要任务上，努力繁荣国内开源生态。大力发展国内开源基金会等开源组织，完善开源软件治理规则，普及开源软件文化。加快建设开源代码托管平台等基础设施。面向重点领域布局开源项目，建设开源社区，汇聚优秀开源人才，构建开源软件生态。加强与国际开源组织交流合作，提升国内企业在全球开源体系中的影响力。

## 欧洲

- 2021 年 9 月, 欧盟委员会 (European Commission) 发布名为《[欧盟经济中开源软硬件对技术独立、竞争力和创新的影响研究](#)》报告, 旨在从多角度研究开源软件 (OSS) 及开源硬件 (OSH) 对欧盟经济的影响, 报告还提供了当前开源软硬件的商业用途、成本、效益等方面的数据, 评估了欧盟通过使用开源软硬件在经济增长、竞争力、增加就业等方面能够达到的潜力。
- [2021 年 12 月 8 日, 欧盟委员会 \(European Commission\) 宣布, 其正在采纳有关开源软件的新规则, 以使之能够在开源许可下发布软件](#)。该规则中指出, 只要对民众、公司或其他社会公共服务有潜在益处, 就可以公开访问其软件解决方案。据悉, 根据新规则, 委员会将通过开源其软件解决方案为公司、初创企业、创新者、公共行政部门等带来重大价值, 这一决定也将刺激创新。

## 二、开源法务合规趋势：意识增强，道阻且长

### 国内

- “完善开源知识产权和法律体系”相继被写入《“十四五”规划和二〇三五年远景目标纲要》、《知识产权强国建设纲要（2021 - 2035 年）》、《“十四五”国家知识产权保护和运用规划》等国家政策文件。
- 浙江省发布全国首个《开源社区知识产权管理规则指引（试行）》，共十二条，包括管理原则、管理平台、社区组成、约束机制、协同研发、软件管理、专利管理、商标管理、风险管理等方面，提出了合法正当、创新引领、应用先导、发展优先以及数据安全的管理原则。
- 深圳中院判决国内首个明确 GPL-3.0 协议法律性质的案例，判定 GPL3.0 协议是一种民事法律行为，具有合同性质，可认定为授权人与用户间订立的著作权协议，属于我国《合同法》调整的范围。
- Elastic License 2.0 (ELv2) 虽然满足源码可见，但并不符合 OSI 的开源定义，采用 ELv2 的 StarRocks 因自称开源，而引起业界非议。
- 国外一开发者在网上抱怨深圳手机公司 UMIDIGI 未公开其采用了 GPL-2.0 协议的手机操作系统内核源代码，位于深圳的知名博主机械妖姬（Naomi Wu）帮助开发者前往手机公司登门索要源代码，并将整个过程的自拍视频发布在了网上，造成热议。事后，该公司在官网的论坛和 GitHub 上相继提供了源代码，国外开发者表达了谢意，纠纷顺利解决。

### 国外

- 近年来，云厂商纷纷将开源软件集成到自己的云产品中，以整体的云服务解决方案推向客户，利益平衡模式面临新的考验。Elastic 公司宣布将 Elasticsearch 和 Kibana 的部分源代码所遵循的协议，由 Apache 2.0 变更为 Server Side Public License (SSPL) 与 Elastic License 双许可，并表示意在抵制云服务提供商，在社区引发巨大争议。SSPL 是由 MongoDB 制定的许可证，含有条款“如果

将程序的功能或修改后的版本作为服务提供给第三方，那么必须免费公开提供服务源代码” Elastic License 要求” 如果将产品作为 SaaS 使用则需要获得商业授权”；两者均未获得 OSI 认证。

- 历经十余年诉讼，美国最高法院在甲骨文诉谷歌版权侵权一案中推翻了联邦巡回法院的判决，裁定谷歌的安卓移动操作系统使用甲骨文 Java API 源代码，属于“合理使用”，不构成版权侵权。

### 三、开源治理成为显学

#### 开源软件供应链治理日趋重要

- 国内企业开始重视并纷纷成立开源计划办公室 (OSPO)
- 安全事件频发，最新爆出的 Apache Log4j2 漏洞，威胁大半个互联网圈（详见下表）

2021 年开源软件供应链安全重大事件表（引用自奇安信集团公众号，[原文链接](#)）

| 序号 | 时间          | 安全事件                                                                            | 成因类型     | 危害                                                                                       |
|----|-------------|---------------------------------------------------------------------------------|----------|------------------------------------------------------------------------------------------|
| 1  | 2020 年 12 月 | 全球著名的网络安全管理<br>软件公司 SolarWinds 遭<br>遇国家级 APT 组织高度复<br>杂的供应链攻击                   | 软件更新包后门  | 导致包括美国关键基础设施、军<br>队、政府等在内的超过 18000 家<br>客户全部收到影响，可任由攻击<br>者操控                            |
| 2  | 2021 年 2 月  | 安全研究任由通过利用开<br>源生态安全机制上的漏洞，<br>实施依赖混淆攻击                                         | 开源生态机制漏洞 | 成功侵入了微软、苹果、PayPal、<br>特斯拉、优步等 35 家国际大型科<br>技公司的内网                                        |
| 3  | 2021 年 3 月  | 攻击者向 git.php.net 服务器<br>上的 php-src 存储库推送了<br>两次恶意提交，在 PHP 代<br>码中植入了一个后门        | 开源库漏洞    | 可通过后面获得运行 PHP 的网站<br>系统的远程代码执行权限                                                         |
| 4  | 2021 年 4 月  | 知名代码测试公司<br>Codecov 宣布其产品的<br>bash uploader 脚本被攻<br>击者修改                        | 软件产品构建问题 | 用户在使用 Codecov 产品时，<br>会向攻击者的服务器发送敏感信<br>息，可造成软件源代码等机密信<br>息泄露                            |
| 5  | 2021 年 5 月  | 在流行的 Visual Studio<br>Code 扩展中发现严重安全<br>缺陷                                      | 开发环境漏洞   | 可使攻击者危急本地机器，或通<br>过开发人员 IDE 构建和部署系统，<br>这些扩展的下载量超过 200 万                                 |
| 6  | 2021 年 6 月  | 研究人员披露了影响 Linux<br>平台基于 Pling 的自由和开<br>源软件（FOSS）市场的<br>漏洞                       | 软件发布源漏洞  | 攻击者可利用该漏洞进行供应链<br>攻击 XSS 蠕虫并实现远程代码执<br>行（RCE）                                            |
| 7  | 2021 年 7 月  | 攻击者获得 Kaseya 公司<br>后端设施访问权限，在运<br>行于客户现场的安全事件<br>响应工具 VSA 服务器上部<br>署 REvil 勒索软件 | 软件产品自身漏洞 | 通过 VSA 服务器将勒索软件安装<br>到联网工作站，从而感染其它第<br>三方企业网络。攻击发生前，互<br>联网上处于联网状态的 VSA 服<br>务器超过 2200 台 |

|    |             |                                                    |        |                                                                                                                                           |
|----|-------------|----------------------------------------------------|--------|-------------------------------------------------------------------------------------------------------------------------------------------|
| 8  | 2021 年 8 月  | 台湾芯片设计厂商 Realtek 称，其 WiFi 模块的三款开发包（SDK）中存在 4 个严重漏洞 | 开发环境漏洞 | 攻击者可利用这些漏洞攻陷目标设备并以最高权限执行任意代码。SDK 用于至少 65 家厂商制造的近 200 款物联网设备中                                                                              |
| 9  | 2021 年 9 月  | 因使用五年前发布的 RunC v1.0.0-rc2，微软 Azure 容器服务器爆出跨账户接管漏洞  | 开源软件漏洞 | 攻击者可攻陷托管 ACI 的多租户 K8S 集群，接管平台上的其他客户的容器，在其中执行代码并访问部署在平台上的数据                                                                                |
| 10 | 2021 年 12 月 | <a href="#">Apache 开源项目 Log4j2 的远程代码执行漏洞细节被公开</a>  | 开源软件漏洞 | 可能的受影响应用包括但不限于：Spring-Boot-strater-log4j2、Apache Struts2、Apache Solr、Apache Flink、Apache Druid、Elasticsearch、Flume、Redis、Logstash、Kafka 等 |

## 国际标准协作

- Linux 基金会推动开源社区衡量指标 [CHAOSS]
- Linux 基金会积极推动 OpenChain & SPDX [软件供应链治理] 国内外社区接轨与交流
- 2021 年 8 月 17 日，华为加入 OpenChain 项目，并成为董事会成员。
- 中国信通院成为 OpenChain 项目首家国内第三方机构，开展基于《ISO/IEC 5230:2020 OpenChain Specifications》检验，协助企业开展开源合规治理

## 国内标准进展

- 中国信通院建立可信开源标准体系，推动 8 个行标立项
- 中国电子技术标准化研究院推动国家开源标准体系顶层设计，推动相关开源标准立项，分别涵盖术语、元数据、许可证框架、开源项目、开源贡献者、开源治理等方面，首个开源领域国家标准《信息技术 开源 开源许可证框架》获批立项

## 木兰开源社区动向

- 木兰开源社区的木兰宽松许可证已有多达 10 万个国内项目采用
- 木兰开源社区吸纳 OpenDigger 项目进入孵化，推动和完善开源项目、社区衡量指标建设

# 四、国际基金会的左右博弈：RMS 重回自由软件基金会与 Rust 社区争议

2021 年 3 月 21 日，自由软件之父理查德·斯托曼（Richard M. Stallman）在 LibrePlanet 2021 年度会议上公开宣布，他重返自由软件基金会（Free Software Foundation），并再次成为基金会的董事

会成员。该消息引发自由及开源软件领域的激烈争论，许多人对斯托曼的回归表示欢迎，但也有众多个人和组织表示反对：开放源代码促进会（Open Source Initiative）呼吁将斯托曼逐出自由软件基金会的董事会，否则将停止与自由软件基金会的合作；红帽（Red Hat）也发表声明称，将暂停对自由软件基金会的所有资助；数千名个人和组织联名发布抵制信，要求罢免斯托曼。据悉，斯托曼在 2019 年因发表不当言论而辞去自由软件基金会的领导职务；反对者声称斯托曼“长期以来一直是自由软件社区的一股危险力量”，斯托曼被指控“歧视女性、残疾人和跨性别者”。

作为自由软件运动的标志性人物，理查德·斯托曼发起 GNU 项目并成立了自由软件基金会。自上世纪八十年代以来，斯托曼一直是重要的自由软件活动家。斯托曼回归引发的系列争议反映出自由和开源软件社区目前存在的理念分歧。自由软件运动和相关的开放源代码运动经历数十年的发展，早已成为一股不容忽视的潮流。然而伴随自由和开源软件社区的壮大，不同群体的理念差异也显现端倪，斯托曼的争议言论使其备受女权主义者等群体的批评。

2021 年的另一场风波同样暴露出自由和开源软件社区面临的复杂态势。Rust 编程语言项目的审核团队 Rust Moderation Team 于 11 月 23 日在 GitHub 上发表[辞职](#)公告，抗议 Rust 核心团队“除了自己，不受任何人的约束”。虽然辞呈未透露过多细节，但原审核团队部分成员的后续言论表明，此次集体辞职事件与 Rust 软件项目长期不断恶化的矛盾有关。另外根据长期参与 Rust 社区的 Dragdu，撰文透露，社区内部的分歧还与“Abolish ICE”运动有所关联。尽管 Rust 社区在诞生之初便倡导和推广“Rustacean Principles”原则以凝聚共识，2018 年还成立了专门的“治理工作组”，但社区矛盾未得到有效缓和。虽然 Rust 编程语言项目近年大获成功，但此次集体辞职事件却也反映了 Rust 社区的治理难题。

开源软件项目需要不同主体的通力协作，社区内部的冲突无疑会对软件项目的开发和维护产生负面影响。在社会价值观念差异愈发加剧的背景下，自由及开源软件社区必须面对如何平衡各方诉求的难题。

## 五、中国开源走向世界，塑造新时代影响力

中国开发者在开源世界的影响力正在不断提升。据最新的 GitHub 年度开发者报告，2021 年中国在 GitHub 上的开发者新增了近 103 万，累计约 755 万。在 ASF、LF、CNCF 等国际开源基金会中，源自中国的开源项目越来越多，同时，开始有更多中国开源人当选国际基金会重要职位。

## Linux 基金会项目列表

- CNCF 基金会：目前有 26 个源自中国的开源项目，毕业项目有 2 个，孵化中项目 2 个，沙盒项目 22 个。其中 2021 年捐赠给 CNCF 基金会的开源项目就达到了 14 个。其中除了来自阿里、华为、腾讯等大厂捐赠的项目，也能看到工业届和学术界的合作以及 Seconed State 和灵雀云等创业公司的项目。
- AI & Data 基金会：目前有 6 个项目。

| 项目名称         | 基金会       | 子基金会                              | 项目捐赠单位 / 方            | 沙箱          | 孵化          | 毕业时间       |
|--------------|-----------|-----------------------------------|-----------------------|-------------|-------------|------------|
| Occlum LibOS | Linux 基金会 | Confidential Computing Consortium | 蚂蚁集团                  |             |             |            |
| harbor       | Linux 基金会 | CNCF                              | VMWare 中国研发中心         | 2018 年 8 月  | 2018 年 11 月 | 2020 年 6 月 |
| TiKV         | Linux 基金会 | CNCF                              | PingCAP               | 2018 年 8 月  | 2019 年 5 月  | 2020 年 9 月 |
| Dragonfly    | Linux 基金会 | CNCF                              | 阿里                    | 2018 年 10 月 | 2020 年 4 月  |            |
| KubeEdge     | Linux 基金会 | CNCF                              | 华为                    | 2019 年 3 月  | 2020 年 9 月  |            |
| ChubaoFS     | Linux 基金会 | CNCF                              | 京东                    | 2020 年 1 月  |             |            |
| Volcano      | Linux 基金会 | CNCF                              | 华为                    | 2020 年 4 月  |             |            |
| BFE          | Linux 基金会 | CNCF                              | 百度                    | 2020 年 6 月  |             |            |
| CNI-Genie    | Linux 基金会 | CNCF                              | 华为                    | 2020 年 6 月  |             |            |
| Chaos Mesh   | Linux 基金会 | CNCF                              | PingCAP               | 2020 年 7 月  |             |            |
| K3s          | Linux 基金会 | CNCF                              | Rancher               | 2020 年 8 月  |             |            |
| OpenYurt     | Linux 基金会 | CNCF                              | 阿里                    | 2020 年 9 月  |             |            |
| OpenKruise   | Linux 基金会 | CNCF                              | 阿里                    | 2020 年 11 月 |             |            |
| Kube-OVN     | Linux 基金会 | CNCF                              | 灵雀云                   | 2021 年 1 月  |             |            |
| Fluid        | Linux 基金会 | CNCF                              | 南京大学、阿里云、Alluxio 开源社区 | 2021 年 4 月  |             |            |
| Vineyard     | Linux 基金会 | CNCF                              | 阿里                    | 2021 年 4 月  |             |            |
| ChaosBlade   | Linux 基金会 | CNCF                              | 阿里                    | 2021 年 5 月  |             |            |
| KubeDL       | Linux 基金会 | CNCF                              | 阿里                    | 2021 年 6 月  |             |            |
| KubeVela     | Linux 基金会 | CNCF                              | 阿里                    | 2021 年 6 月  |             |            |
| WasmEdge     | Linux 基金会 | CNCF                              | Second State          | 2021 年 6 月  |             |            |
| Karmada      | Linux 基金会 | CNCF                              | 华为                    | 2021 年 6 月  |             |            |

|                         |           |                        |                                                                          |             |  |  |
|-------------------------|-----------|------------------------|--------------------------------------------------------------------------|-------------|--|--|
| Inclavare Containers    | Linux 基金会 | CNCF                   | 阿里                                                                       | 2021 年 9 月  |  |  |
| SuperEdge               | Linux 基金会 | CNCF                   | 腾讯                                                                       | 2021 年 9 月  |  |  |
| Nocalhost               | Linux 基金会 | CNCF                   | 腾讯                                                                       | 2021 年 11 月 |  |  |
| Open Cluster Management | Linux 基金会 | CNCF                   | 阿里、红帽、蚂蚁                                                                 | 2021 年 11 月 |  |  |
| OpenELB                 | Linux 基金会 | CNCF                   | 青云                                                                       | 2021 年 11 月 |  |  |
| Piraeus-Datastore       | Linux 基金会 | CNCF                   | LINBIT、DaoCloud、上海浦东发展银行                                                 | 2021 年 11 月 |  |  |
| EdgeGallery             | Linux 基金会 | Edge                   | 华为                                                                       | 2021 年 10 月 |  |  |
| Baetyl                  | Linux 基金会 | Edge                   | 百度                                                                       | 2019 年      |  |  |
| EMQ X Kuiper            | Linux 基金会 | Edge                   | EMQ                                                                      | 2021 年 7 月  |  |  |
| FATE                    | Linux 基金会 |                        | 微众银行                                                                     | 2019 年 6 月  |  |  |
| Caliper                 | Linux 基金会 | Hyperledger            | 华为                                                                       | 2018 年 3 月  |  |  |
| cello                   | Linux 基金会 | Hyperledger            | Baohua Yang (Oracle)、Haitao Yue (IBM)、Tong Li (IBM)、Jiahao Chen (VMware) | 2017 年 1 月  |  |  |
| Adlik                   | Linux 基金会 | AI & Data              | 中兴通讯                                                                     | 2019 年 10 月 |  |  |
| Angel                   | Linux 基金会 | AI & Data              | 腾讯                                                                       | 2021 年 5 月  |  |  |
| DELTA                   | Linux 基金会 | AI & Data              | 滴滴                                                                       | 2019 年 10 月 |  |  |
| Elastic Deep Learning   | Linux 基金会 | AI & Data              | 百度                                                                       | 2020 年 7 月  |  |  |
| Milvus                  | Linux 基金会 | AI & Data              | Zilliz                                                                   | 2019 年 12 月 |  |  |
| OpenBytes               | Linux 基金会 | AI & Data              | Graviti                                                                  | 2021 年 11 月 |  |  |
| ONAP                    | Linux 基金会 | Networking             | 华为、中兴                                                                    | 2018 年 1 月  |  |  |
| XGVela                  | Linux 基金会 | Networking             | 中国移动                                                                     | 2020 年 5 月  |  |  |
| Feilong                 | Linux 基金会 | Open Mainframe Project |                                                                          |             |  |  |

## Linux 子基金会列表

| 子基金会                | 发起单位       | 发起时间        |
|---------------------|------------|-------------|
| OpenMessaging       | 阿里         | 2017 年 10 月 |
| NextArch Foundation | 40 家企业联合发起 | 2021 年 11 月 |
| SODA Foundation     | 华为         | 2020 年 6 月  |
| TARS Foundation     | 10 家企业联合发起 | 2020 年 3 月  |

## Apache 软件基金会项目列表

- Apache 软件基金会: 目前有 24 个源自中国的开源项目, 其中顶级项目达到了 14 个, 孵化项目 10 个。而 2021 年 ASF 孵化器新增的 5 个项目全部来自中国。<https://projects.apache.org/projects.html>

| 项目名称                    | 基金会          | 项目捐赠单位 / 方   | 孵化         | 毕业时间       |
|-------------------------|--------------|--------------|------------|------------|
| Apache DolphinScheduler | Apache 软件基金会 | 易观           | 2021/4/8   | 2021/4/8   |
| Apache ECharts          | Apache 软件基金会 | 百度           | 2018/1/18  | 2020/12/16 |
| Apache Ozone            | Apache 软件基金会 | 原 Hadoop 子项目 | 2018/11/22 | 2020/10/21 |
| Apache IoTDB            | Apache 软件基金会 | 清华大学         | 2018/11/18 | 2020/9/17  |
| Apache APISIX           | Apache 软件基金会 | 深圳支流科技       | 2019/10/17 | 2020/7/15  |
| Apache ShardingSphere   | Apache 软件基金会 | 京东科技         | 2018/11/10 | 2020/4/16  |
| Apache Dubbo            | Apache 软件基金会 | 阿里巴巴         | 2018/2/16  | 2019/5/15  |
| Apache SkyWalking       | Apache 软件基金会 | 吴晟           | 2017/12/8  | 2019/4/17  |
| Apache Griffin          | Apache 软件基金会 | eBay         | 2016/12/5  | 2018/11/21 |
| Apache ServiceComb      | Apache 软件基金会 | 华为           | 2017/11/22 | 2018/10/17 |
| Apache HAWQ             | Apache 软件基金会 | Pivotal      | 2015/9/4   | 2018/8/15  |
| Apache RocketMQ         | Apache 软件基金会 | 阿里巴巴         | 2016/11/21 | 2017/9/20  |
| Apache CarbonData       | Apache 软件基金会 | 华为           | 2016/6/3   | 2017/4/19  |
| Apache Kylin            | Apache 软件基金会 | eBay         | 2014/11/25 | 2015/11/18 |
| Apache SeaTunnel        | Apache 软件基金会 | 乐视           | 2021/12/9  |            |
| Apache Linkis           | Apache 软件基金会 | 微众银行         | 2021/8/2   |            |
| Apache Kyubi            | Apache 软件基金会 | 网易           | 2021/6/21  |            |
| Apache ShenYu           | Apache 软件基金会 | Dromara 开源社区 | 2021/5/3   |            |
| Apache EventMesh        | Apache 软件基金会 | 微众银行         | 2021/2/18  |            |
| Apache Pegasus          | Apache 软件基金会 | 小米           | 2020/6/22  |            |
| Apache InLong           | Apache 软件基金会 | 腾讯           | 2019/11/3  |            |
| Apache Teaclave         | Apache 软件基金会 | 百度           | 2019/8/19  |            |
| Apache brpc             | Apache 软件基金会 | 百度           | 2018/11/13 |            |
| Apache Doris            | Apache 软件基金会 | 百度           | 2018/7/18  |            |

## 中国开源人当选国际基金会重要职位

- 吴晟：当选 Apache 软件基金会董事
- 堵俊平：当选 LF AI & Data 基金会董事会主席
- 王晔晖：当选 Linux 基金会 CHAOSS 项目的董事会的董事
- 单致豪：当选 TARS 基金会董事主席，FinOps 基金会董事，LF Research 董事

## 六、开源新创投资持续发光发热

2020 年底到 2021 底，国内外基于开源项目的初创企业空前活跃，基于开源项目的商业公司获得融资、上市等现象屡见不鲜，且融资金额和估值 / 市值不断刷新上限，开源社区和开源软件的商业价值得到资本认可。以下为部分开源新创公司的融资和上市情况：

### 国外

- 6 月，开源产品分析工具 PostHog 融资 1,500 万美元
- [GitLab 在纳斯达克股票交易所上市](#) ( 2021-10-15 )
- 红帽峰会发布数据分析类产品套件
- Eclipse 经历大裁员，无力运营将项目捐献给 Linux 基金会
- Databricks 估值 380 亿美金，计划 2022 年 IPO
- SUSE 在法兰克福证券交易所上市
- Docker 融资 2,300 万美元，与微软合作转型开发工具公司
- 11 月 11 日，开源工具 CloudQuery 获得 350 万美元种子轮融资
- 11 月 30 日，Upbound 投资 6,000 万美元发展开源 Crossplane 多云管理项目
- 12 月 9 日，聚焦云基础设施的软件企业 [Hashicorp](#) 在美国纳斯达克上市，市值超过 150 亿美元
- 12 月 16 日，Cockroach Labs [再获 2.73 亿美元 F 轮融资](#)，估值已经达到 50 亿美元，而今年 1 月 Cockroach Labs 才刚获得了 1.6 亿美元融资

### 国内

- 白鲸开源 ( 基于 Apache DolphinScheduler ) 获得数百万美元天使轮融资
- 北京思斐软件 SphereEx ( 基于 Apache ShardingSphere 项目 ) 完成数百万美元天使轮融资
- Engula 天使轮融资

- Authing 完成 500 万美金 Pre-A 轮融资以及 A 轮融资 2,300 万美元
- 深圳支流科技 API7 ( 基于 Apache APISIX 项目 ) 完成数百万美金 Pre-A 轮以及 A 轮融资
- StreamNative ( 基于 Apache Pulsar 项目 ) 完成 2,300 万美元 A 轮融资
- 极纳科技 Jina AI 完成 3,000 万美元 A 轮融资
- 一流科技 OneFlow 完成 5,000 万人民币 A 轮融资
- 上海赓睿信息科技 Zilliz 完成 4,300 万美元 B 轮融资
- 涛思数据 TaosData 完成 4,700 万美元 B 轮融资
- 上海跬智信息技术 Kyligence ( 基于 Apache Kylin 项目 ) 完成 7,000 万美金 D 轮融资
- PingCAP 已完成 5 轮融资，最新一轮是去年 11 月 2.7 亿美金 D 轮融资，融资额已达到 3.4 亿美元，而估值约 30 亿美元，为国内开源新创公司的现象级里程碑
- 鲸鯨科技获创新工场领投 1000 万美元天使轮融资

## 七、开源操作系统迎来新一轮繁荣期

2020 年底，Red Hat 宣布将在 2021 年底结束对 CentOS 8 的支持，这对全球操作系统的发展都带来了巨大的影响，全球用户开始尝试寻找合适的系统进行切换。解决用户在 CentOS 退出后可能面临的风险，成为了操作系统厂商和研发人员努力的方向。

也是在这样的背景下，2020 年 9 月，阿里云联合统信、龙芯、中科方德等 16 家操作系统、芯片、云计算公司共同发起 OpenAnolis 龙蜥操作系统开源社区。龙蜥操作系统源于 2011 年前阿里为替换自身的 CentOS 而打造的 Alibaba Cloud Linux，经历了“双 11”百万虚拟机、千万部署的实际考验。2021 年 11 月 4 日，龙蜥操作系统宣布将捐赠到开放原子开源基金会进行孵化。目前，龙蜥操作系统已在阿里云全面上线，总装机量达百万量级。龙蜥社区（OpenAnolis）已拥有 50 多家生态企业。其中统信软件、中国移动云等已基于龙蜥操作系统发布商业版本。龙蜥也是国内首个从操作系统层面提供全软件栈国密算法的 OS 解决方案，并且对内核 SM4 算法做了深度优化，性能提升近 800%，让中国国密算法从合规走向生产应用。

在更早之前的 2019 年 12 月 31 日，华为将内部自研且已经过十年打磨的操作系统 EulerOS 重新命名为 openEuler 正式开源，源代码、镜像及开发测试环境全部向社区开放。开源近两年，openEuler 共

发布了四个版本，分别是 20.03 LTS 版本、20.09 创新版本、21.03 创新版本和 21.09 创新版本。2021 年 3 月，openEuler 在内核热升级和内存分级管理上做了创新，21.03 创新版本推出；9 月，[21.09 创新版本发布](#)，该版本不仅增强了服务器和云计算场景能力，还实现了对于边缘计算和嵌入式场景的支持。

2021 年 11 月，华为 openEuler 正式捐赠给开放原子开源基金会。随后，英特尔也正式签署贡献者许可协议加入 openEuler 开源社区。如今 openEuler 社区已经有 300 多家企业、近万名社区开发者加入。

2020 年 9 月，腾讯也将 TencentOS Tiny 捐献给了开放原子开源基金会。目前开放原子开源基金会已在孵化 5 个国产开源操作系统项目，分别是龙蜥操作系统（Anolis OS）、openEuler、OpenHarmony、TencentOS Tiny、AliOS Things，此外还有类 Redis 存储系统 PIKA、云原生分布式数据库 ZNBase 等，隐有成为中国开源基础软件大本营之势。

2021 年 10 月 28 日，开放原子开源基金会技术监督委员会投票通过开源项目 OpenCloudOS 进入开放原子开源基金会孵化。12 月 22 日，国产开源操作系统 OpenCloudOS 的开源社区正式成立。腾讯、宝德、北京初心、北京红旗、飞腾、浪潮、龙芯中科、OPPO、先进开源、中电科申泰、中科方德、兆芯等 20 余家操作系统生态厂商及用户成为首批创始单位。据悉，作为国产开源操作系统社区，OpenCloudOS 沉淀了多家参与单位在软件和开源生态的优势，在云原生、稳定性、性能、硬件支持等方面均有坚实支撑，可以平等全面地支持所有硬件平台。

在移动开源操作系统方面，2021 年 1 月 31 日，JingOS 开放下载，鲸鯨科技的 JingOS 基于 Linux 内核，是一款主要适用于平板电脑，同时也可支持笔记本、手机等多种终端的新一代通用型移动操作系统，开源代码贡献量超 115 万行，三次登陆 YC 的 Hacker News 头条，是 2021 年上半年全球开源社区关注度最高的移动 Linux 操作系统。

## 八、Rust 迈上新征程

作为一门开源的通用系统级编程语言，Rust 由于其出色的内存安全机制、不亚于 C 语言的性能优势等特点，吸引了大量开发人员关注。2021 年，Rust 告别“动荡”，在多个维度上都迈出了一大步。

### 专家点评

段夕华：在基础软件及隐私计算领域领域，Rust 必将大有所为。

2021 年 2 月 9 日, [Rust 基金会](#)由 Mozilla、Amazon、华为、谷歌、微软联合[成立](#)。作为创始白金成员, 五家企业承诺在两年时间里, 每年投入不少于 100 万美元的预算, 用于 Rust 项目的开发、维护和推广。Rust 基金会的成立, 一方面让 Rust 技术的研发和推广免去资金之忧, 另一方面, 多方参与的基金会管理模式也可以避免单一企业对社区的垄断, 确保 Rust 开源社区的开放性与多元化。

Rust 基金会成立前后, 硅谷同步掀起了一场 [Rust 人才争夺大战](#)。不少原 Rust 团队的活跃开发人员纷纷转而加入谷歌、微软、Amazon、Facebook 等科技巨头, 这也成为 Rust 语言未来看涨的明确信号。首先, Rust 开发者被商业公司招聘并继续在社区专职投入, 会是对社区与技术可持续发展非常好的保证; 而这些开发者在解决商业公司遇到的技术问题的同时, 反过来将改进方案贡献回社区, 更能增加 Rust 在大规模商用场景下的效率与技术韧性, 结果将是双赢的。

2021 年 11 月, ARM、Sentry、Knóldus、Spectral、Automata、Activision 和 Toyota Connected、Clever Cloud、Ferrous Systems、Futurewei、KDAB、Open Source Security、ParaState、Tag1、Zama 共 [15 家企业也相继加入 Rust 基金会](#)。

除了商业公司大力支持, Linux 社区也对 Rust 表现出了积极的接纳态度:

- 2021 年 6 月, [谷歌强推 Rust 进驻 Linux 内核](#), 与 [Rust for Linux 项目](#)的主要开发者兼 Linux 内核开发者 Miguel Ojeda 签订合同;
- [2021 年 7 月 4 日](#), Linux 内核团队发布添加 Rust 支持的“v1”补丁;
- [2021 年 12 月 6 日](#), Linux 内核团队发布支持 Rust 的“v2”补丁。

Miguel Ojeda 在邮件中指出, 新的“v2”补丁将为 Linux 内核增加对 [Rust 作为第二语言的支持](#), 并且对 Rust 的整体支持进行了多项改进。目前 Rust for Linux 项目进展顺利, 开发者有望在 2022 年看到 Linux 内核正式支持 Rust。

## 九、AI & 低代码将会如何改变开源, 值得关注

- 2021 年 3 月, 腾讯云微搭 WeDa 低代码平台正式发布
- 2021 年 3 月, 微软宣布推出新的开源编程语言 [Power Fx](#), 一种基于 Microsoft Excel 的低代码公式语言; 将可以在整个 Microsoft Power Platform 中进行使用。

- 2021 年 6 月，GitHub 推出 Copilot，一款由 GitHub 和 OpenAI 开发的人工智能工具，通过自动完成代码来帮助 Visual Studio Code、Neovim 和 JetBrains 的用户。GitHub Copilot 发布不足一周时间，却引发了前所未有的争议。有开发者表示：「我不同意 GitHub 在未经授权和未经许可的情况下，使用受版权保护的源代码作为其 Copilot 产品的训练数据。该产品将受版权保护的源代码放入使用者的软件中，而不告知他们源代码的许可，这导致了对版权所有者作品的未经授权和未经许可的不当使用。」

## 专家点评

段夕华：除了合规挑战之外，目下基于 ML 的自动代码生成及补足，因为其学习样本存在着质量差异且未做深入甄别，其质量也会是个很大的问题，但定位为辅助代码工具或将有所作为。

- 2021 年 7 月，OpenAI 宣布推出一种类似于 Python 的开源编程语言 [Triton](#)，并发布了 Triton 1.0 版本。目标是替代 Nvidia CUDA，可使没有 CUDA 经验的研究人员能够编写高效的 GPU 代码，且大多数情况下可以与专家所能编写的代码质量媲美。
- 2021 年 11 月，已经有文章报道《[20 个开源低代码平台](#)》，并总结了低代码平台的六大好处。
- 2021 年 12 月，微软专家推出新型 AI 工具，能自动发现代码中的 Bug。据 mspoweruser 报道，微软公司的研究人员开发了一种新的人工智能（AI）工具，利用深度学习算法可以发现代码中的错误，帮助开发者更准确、更高效地对程序进行调试。这种工具能够识别一些常见的错误，比如不正确的符号“<=”和“<”，不正确的布尔运算符“and”和“or”，变量滥用等等。为了测试该系统，微软使用了 Python 代码。与其它方法相比，经过训练后的模型发现 Bug 的概率最多可以提高 30%。【[来源](#)】

一个值得思考的问题是：当开源不断的推动各种各样的技术飞速向前发展时，这些技术将会如何反过来促进开源技术、开源开发方式的发展？

## 十、开源硬件持续升温，RISC-V 成果涌现

- 2021 年 1 月，欧洲核子研究组织推出的 CERN 开放硬件许可证 2.0 版（CERN Open Hardware Licence Version 2）获得 OSI 认证，包括 Permissive、Weakly Reciprocal、Strongly Reciprocal 三个版本。
- 2021 年 6 月 21-26 日，首届 RISC-V 中国峰会（RISC-V World Conference China）在上海科技大学举办，峰会由上海科技大学和中国科学院软件研究所联合主办。
- 2021 年 6 月 22 日，中科院计算所牵头开发的香山开源高性能 RISC-V 处理器在首届 RISC-V 中国峰会发布；7 月，香山处理器完成第一版（代号雁栖湖）的 28nm 流片；10 月 30 日，香山处理器

团队负责人包云岗研究员在第六届中国开源年会发表演讲《开源芯片的挑战与机遇》；12月6日，香山通过 RISC-V Summit 第一次在国际 RISC-V 社区正式亮相。

- 2021年9月，RISC-V 国际协会、CHIPS 联盟、OpenPOWER 基金会和西部数据公司联合筹建开放硬件多样性联盟（Open Hardware Diversity Alliance），努力推动为开放硬件社区中的女性和未被充分代表的个体提供支持。
- 2021年10月19日，阿里平头哥于 2021 云栖大会宣布开源四款玄铁 RISC-V 系列处理器，成为系列处理器与基础软件的全球首次全栈开源。
- 2021年10月28日，OpenPOWER Summit NA 2021 在线上举办，OpenPOWER 基金会执行董事 James Kulina 在开场演讲中提到，采用开源硬件是大势所趋。
- 2021年12月6日，Linux 基金会 2021 年度报告发布，报告体现出对开源硬件的重点关注：Linux 基金会社区关注的方向包括开源硬件、开放标准、开放数据、开放硬件，分别占比 75%、20%、3%、2%；Linux 基金会将会在下一阶段创建新的开放硬件社区，例如和 Open19 创建开放硬件生态。
- 2021年12月9日，由赛昉科技支持的 RISC-V 开源社区平台 RVspace 正式上线。

---

## 感谢

---



思否



主办：开源社

协助：云启资本、Gitee、infoQ、ALC Beijing、思否、  
CHAOSS、X-lab，木兰开源社区

---

## 参与志愿者

---

召集人：王伟

问卷篇：王皓月

数据篇（Github）：顾业鸣、赵生宇

数据篇（Gitee）：李泽辰

商业化篇：云启资本

开源大事记：阿法兔、蔡芳芳、高原、耿航、郭雪、江炜婕、李圳虎、  
梁尧、刘天栋、王蓉、卫剑钊、薛亮、杨丽蕴、张俊霞、  
庄表伟

整体报告汇总 / 编辑 / 设计 / 排版：李明康、朱庆裕、朱亿钦、王蓉

特别感谢开源社 – 顾问委员会的点评和参与！