

开源如何可能？

从知识生产与传播的角度分析

庄表伟 2024-11-16

提纲

什么是开源？

开源的起源与发展

知识生产与传播的演变历程

开源与知识生产和传播的关系

开源世界的知识图谱

当今开源面临的挑战

结论：开源的价值与启示

什么是开源

关于开源的3种定义

开源的定义——面向学生版

- 一个学生要完成一个作业。
 - 他做出了一个初步的成果。
 - 他没有直接交给老师，而是邀请所有的同学：大家不要各自做了，一起来做一个吧。
 - 全班所有的同学，都基于他的初步成果讨论、改进。
 - 最后提交给老师，一个更加完善的作业。
- 将这个做作业的过程映射一下：
 - 放大到整个世界，而不再局限于一个班级
 - 在互联网上进行
 - 作品以数字化形式存在（软件、文档、其他设计）
 - 不是交作业，而是不断发布新版本
- 这基本上就算是开源了

开源的定义——朴素版

- 开源——Open Source。
- 源（源代码）——Source Code，通常是指计算机软件的源代码。
- 开（开放）——Open，首先是指开放可获取（Open Access），所有人都能够获得软件的源代码。
- 开放的源代码——可以被任何人查看、使用、修改和分发的源代码。
 - 与私有软件不同，开源软件的源代码是公开的，任何人都可以参与其中并做出贡献。
 - 这种模式具有很多好处，例如可以促进创新、降低开发成本、提高软件质量等等。
- 开源也可以作为一个形容词
 - 除了开源软件，还有开源硬件、开放文档、开放数据等等。这些衍生事物也具备开源软件的很多特点
- 开源作为一个动词
 - 例如，将某个软件开源，就意味着它的源代码可以被任何人查看和修改。这样，就可以吸引更多的开发者来为它做出贡献，从而让它变得更加完善和稳定。

开源的定义——OSD版

- OSD（The Open Source Definition，开源软件定义），诞生于1998年，长期由OSI（Open Source Initiative）维护
- OSD的开头：Open source doesn't just mean access to the source code。
 - 开源不仅仅是指能够访问源代码。
 - 开源软件的分发条款必须符合以下十条标准：
 1. 自由再分发
 2. 源代码
 3. 派生作品
 4. 作者源代码的完整性
 5. 不歧视个人或团体
 6. 不歧视领域
 7. 许可证的分发
 8. 许可证不得特定于产品
 9. 许可证不得限制其他软件
 10. 许可证必须技术中立

开源的起源与发展

从早期黑客到现代开源运动

黑客文化与黑客伦理

黑客文化起源于1960年代末的大学校园、俱乐部与科研机构，如MIT人工智能实验室。这些地方形成了独特的技术风格和文化，其中包括Jargon File，这是一份由黑客自己所建立和维护的手册，记录了黑客传统、民间传说和幽默。黑客被视为技术领域的极客和创新者，他们对技术抱有深厚的热情，并推动技术的极限。

黑客伦理的核心原则：

- 信息自由：信息应该是自由流通的，限制信息的流通是对知识自由的损害。
- 对计算机的渴望：黑客对计算机和技术的无限渴望推动他们深入探索和创造。
- 反对权威：对权威的怀疑和对自主学习的强调，鼓励个人自由和创新。
- 改善世界：黑客相信他们的技能有责任用于改善社会和世界。

自由软件运动时间线

1969年：Unix的诞生。Unix的代码在早期被Unix社区的Hacker们自由共享，对后来的自由和开源软件运动产生了重要影响。

1983年9月：Richard Stallman 在 Usenet 上宣布了 GNU 项目，旨在创建一个“自由的Unix”。

1989年：GPLv1 发布

2001年：成立自由软件基金会欧洲分部，支持自由软件并反对欧洲的软件专利。

1976年：Richard Stallman 开发了最初的 Emacs，后来也发布了遵守GNU通用公共许可证的GNU Emacs。

1985年：Stallman 成立了自由软件基金会（FSF）。

1991年：GPLv2 发布。Linus Torvalds 启动了 Linux 内核项目，这个项目后来也采用了 GPLv2 。

2007年：GPLv3 发布。FSF启动了AGPL的开发。

自由软件的定义与GPL

自由软件的定义

- 在自由软件哲学中，通常认为自由软件应该遵循四个基本标准。这些标准如下：
 - 依照你的想法执行该程式的自由，无论任何目的（自由之零）
 - 研究该程式如何运作的自由，并依照你的想法修改它以符合你的运算所需（自由之壹）。
 - 能存取程式的源始码 (source code) 是这项自由的先决条件。再次散布程式副本的自由，如此你就能帮助他人（自由之贰）。
 - 将你修改过后的版本散布给他人的自由，如此你就有机会让你的改善惠及社群整体。能存取源始码是这项自由的先决条件（自由之参）
- 自由软件的目标是为用户提供自由和控制力，以使其能够自由地使用、复制、分发和改进软件，而不受专有软件的限制。自由软件的哲学是自由软件运动的核心，自由软件基金会是该运动的主要推动者之一。



GNU通用公共许可证（GPL）

- GPL是自由软件基金会推广的一种自由软件许可证。它确保了用户对软件的自由使用权，包括自由地运行、研究、复制、修改和分发软件。GPL的核心原则是要求任何基于GPL许可的软件必须以相同的自由许可证进行发布，以确保软件的自由性得以延续。

Linus其人与Linux的早期版本

Linus Torvalds，芬兰裔美国软件工程师，以创建和主导Linux内核开发而闻名。在赫尔辛基大学期间，因对计算机的浓厚兴趣，开发了Linux。

- 在Linus的传记中提到，他对计算机的兴趣始于童年。他的爷爷是一名统计学教授，家里有一台早期的计算机，Linus就是在那台机器上开始探索编程的。

Linux的早期版本由Linus Torvalds在1991年发布，标志着开源操作系统的诞生。最初仅为个人项目，但很快吸引了全球开发者的贡献和兴趣。

Linux凭借其稳定性、安全性和开源特性，成为了多种计算平台的重要选择，极大地推动了开源运动的发展。

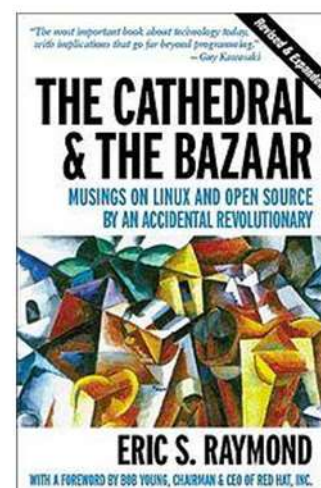


Eric Raymond的《大教堂与集市》

Eric Steven Raymond（通常被称为 ESR），美国软件开发人员、开源软件倡导者，1997 年发表的论文和 1999 年出版的《大教堂与集市》一书的作者。他为 Roguelike 游戏《NetHack》编写了一本指导手册。20 世纪 90 年代，他编辑并更新了《术语文件》，并作为《新黑客词典》出版。

《大教堂与集市》是Eric Raymond的著作，其中探讨了软件开发的两种模式：类似于密闭、中心化的大教堂模式，和开放、协作的集市模式。Raymond通过Linux内核开发的例子，阐明了集市模式下，如何通过广泛协作和开源的方式，有效地提升软件质量和创新速度。这本书对开源软件运动产生了深远的影响。

- 礼物文化 & Linux眼球定理
- 创建优秀开源软件的19条经验



Firefox开源与Open Source的诞生

前情提要

- 1997年，Netscape公司高管，正好听到了Eric Raymond的演讲
- Netscape邀请Eric Raymond共商开源大计

1998年1月，Netscape对外宣布：将在本季度内对外开源

1998年2月，Netscape浏览器源代码对外开源

1998年2月3日，在美国加利福尼亚帕罗奥多举办的一次战略会议上，与会者达成一致意见，“Open Source”这个名词诞生

1998年2月，开放源代码促进会（Open Source Initiative，缩写：OSI）正式成立

1998年4月7日，第一届Open Source Summit召开



用开源对抗微软的联盟如何形成？

1990年代，IBM为支持Linux而缴纳了10亿美元的专利费

1999年，IBM帮助创建了Apache软件基金会，提供了上万行代码和资源以支持Apache Web Server Project。

2000年，开源发展实验室（OSDL）成立。

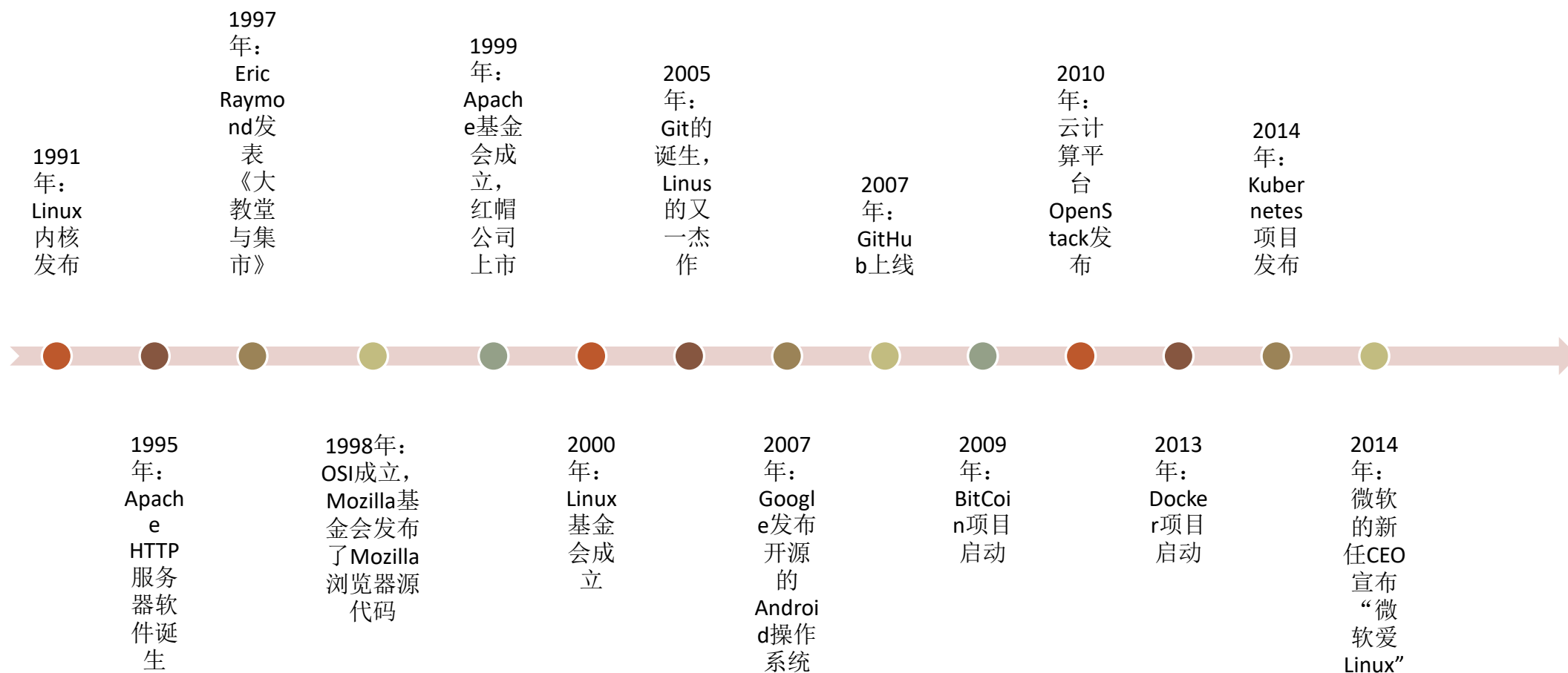
2001年，IBM开发了Eclipse项目并引导了2004年Eclipse基金会的成立

2005年底，IBM 与 Red Hat、SUSE、索尼、NEC 和飞利浦等企业携手，创立了 OIN（Open Invention Network），以寻求解决技术领域知识产权问题的新途径。

2007年，与自由标准组织（FSF）合并后从而成为Linux基金会，目前包括AT&T、思科系统、富士通、日立、华为、IBM、英特尔、**微软(2016年)**、日本电气、甲骨文公司、高通、三星电子等知名的科技公司



开源运动时间线



现代开源运动的本质特征

一、开源项目的本质

- **开源项目**的本质，就是一群人以**开放式协作**的方式，组成一个开源社区，共同创造**数字公共产品**。
- **开放式协作**的本质是：基于规则的开放参与，且规则本身对与全体参与者来说，也是开放可修改的。
- **数字公共产品**的本质是：以数字形式凝聚的人类**知识**，这些知识可以以多种方式，节约人类的时间。

二、开源生态的本质

- **开源生态**的本质，就是由各种各样授权协议/规则所界定的、由全部数字公共产品组成的**数字知识公地**。所有参与这个开源生态的主体，基于共同遵守的规则，形成了开放利益共同体。
- **数字知识公地**的本质：与公地悲剧中所说的物理世界的公地不同，只要数字世界依然存在，只要复制成本始终接近于0，数字世界的公地就不会减少，只会越来越繁荣。
- **开源世界的规则**的本质，就是约束力大小不等的各种共识，这些共识内容，共识者的数量，共识被执行的可能性，共同决定了这个数字知识公地的使用成本。
- **开放利益共同体**的本质，就是各种主体基于各自不同的理由、价值判断，都认为向这个数字知识公地做出贡献，利大于弊。哪怕大家对于利弊的定义和理解，都完全不同。

知识生产与传播

从古代到当代的演变历程

从古代到中世纪

生产：对经典的解读与思考

- 典型的例子是对圣经文本、亚里士多德、柏拉图等哲学家著作的注释工作。

特征：经院哲学体系

- 经院哲学不仅仅是知识生产的体系，更是一种逻辑和思维训练方式。其方法注重辩证法，强调通过辩论解决理论和信仰之间的矛盾。

载体：修道院、手抄书籍和口耳相传

- 修道院不仅是知识的保存和生产场所，同时也是抄写经典、保存书籍的中心。此时期，欧洲的许多知识主要在修道院的手工书写室中被抄录下来。
- 除了手抄书籍和口耳相传，少数流通的抄本也通过集市、商队等途径实现有限的跨区域传播。

传播：知识的权威化与有限传播

- 在中世纪，教会不仅控制了知识传播的内容，也监督了思想的正统性，尤其是通过大学和宗教法庭对“异端”的打击来维持思想的单一性和权威性。



文艺复兴与启蒙运动时期

生产：知识生产以科学探究和理性批判为核心

- 从神学主导的知识体系过渡到科学和理性的知识体系。科学家通过实验和观察探寻自然规律，开启了“科学革命”，思想家则通过哲学反思推动“启蒙运动”。

特征：知识生产趋向多元化和开放化

- 开始摆脱宗教的限制，强调人文精神、科学实证和理性思维，形成对自然、社会和人性的新探索。

载体：印刷术的普及让书籍成为主要的知识载体

- 图书馆和沙龙成为知识交流和保存的中心，科学期刊和百科全书也推动了知识的系统化。
- 包括书籍、期刊、图解手册和演讲，尤其是百科全书为知识的整理和普及提供了重要的工具，推动了知识的更广泛共享和应用。

传播：知识传播向公众开放

- 印刷术极大地扩展了知识的覆盖范围，学术社团和公开辩论成为风气，知识开始流向大众。



工业革命与现代化时代

生产：知识生产开始高度专业化，科学研究逐渐向实验室、研究所集中。

- 知识的生产逐渐团队化，注重科学方法和实验数据，服务于工业和技术革新，推动了科学知识体系的发展。

特征：知识生产体系化和标准化

- 以学科为单位划分研究方向。随着教育体系的发展和工业需求的推动，技术知识的积累和专业人才的培养成为社会重点。

载体：书籍、学术期刊和专利文献成为知识的主要载体

- 知识存储和传播逐渐转向公共图书馆和学术机构。教科书、实验室笔记、研究报告和期刊文献作为科学研究的核心资料保存，知识更加大众化和实用化。
- 专利系统的建立推动了技术创新的保护和传播，期刊和研究报告为科学界提供了可靠的学术交流平台。

传播：知识传播更加制度化和组织化

- 学校、大学和技术培训等机构，学术会议、国际展览等交流方式不断增加。



信息化与数字化时代

生产：以数据和信息技术为核心，依托计算机和互联网进行生产

- 涵盖科学研究、数据信息、软件开发等新领域。协同工具、云计算和大数据分析促进了知识生产的速度和创新性。

特征：知识生产去中心化和全球化，开源和分布式协作成为常态

- 任何有互联网接入的人都可以参与知识的生产和改进，个人、团队、企业和社会共同推动创新。人工智能和自动化工具参与知识的生成和优化。

载体：数字化载体取代传统载体

- GitHub、Wikipedia等平台确保知识的全球化存储和即时共享。
- 电子书、视频、播客等新媒体形式广泛应用，AI算法实现个性化推荐，互动平台提升了知识的易用性和可访问性。
- Copyleft的概念被提出。

传播：依托互联网和社交网络，速度快、覆盖广

- 视频教程、在线课程、开放教育资源（OER）等形式加速了知识的普及。
- 开放式协作，则是知识传播的最新方式。



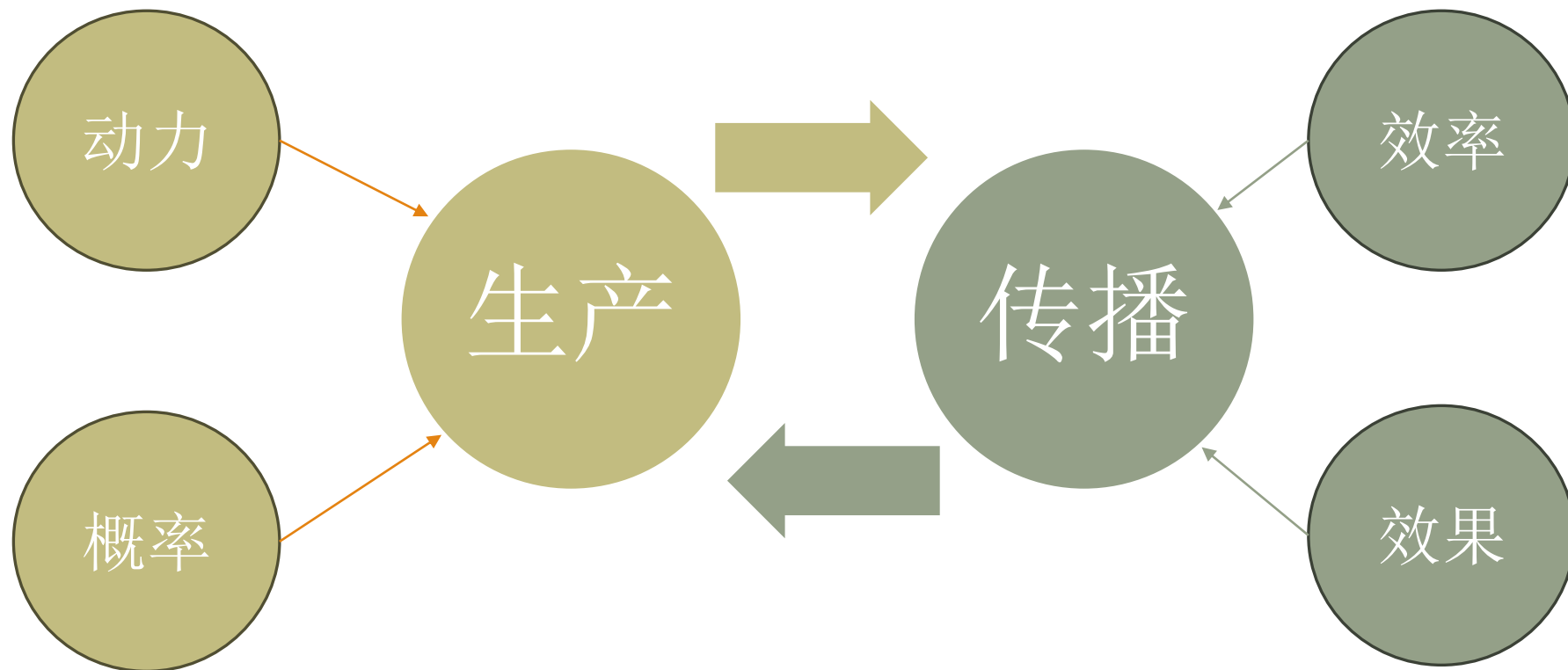
知识生产与传播的激励因素

激励因素：

需求
制度
文化

激励因素：

连通性
活力
世界观



激励因素：

需求
技术
载体

激励因素：

世界观
价值观
全球化程度

开源、知识生产和传播

一些关键性的变化

从开源的角度看知识

什么是知识？

- 知识其实分为很多种，有书本上的，容易普及、传授的知识。也有非常私人化、经验性的知识，只能通过师父带徒弟，口传心授，才能够掌握的知识。
- 书面知识、口头知识、技术、技巧、经验、心得、秘方、专利等等，都是知识的各种形态。
- 开源社区里的代码，就是一种“可以执行”的知识。

什么是知识溢出？

- 维基百科版
 - 在学术界，比较有代表性的定义是哈佛大学的Zvi Griliches（1992）的定义：做相似的工作并从彼此研究中受惠。更加详细的划分又有三种类型：MAR 溢出、Porter 溢出和 Jacobs 溢出。
 - MAR 溢出：同行之间的非竞争性知识溢出
 - Porter 溢出：同行之间的竞争性知识溢出
 - Jacobs 溢出：同城、跨行之间的知识溢出
- 个人版
 - 溢出、转移、扩散、传播，都是从一个人到另一个人，或者从一个企业到另一个企业。有可能是没防住，也可能是无所谓，也可能是主动的传授。
 - 开源社区里流动着各种类型的知识，代码是其中最精准的一种，其他的经验交流、文档与布道，都是开源社区里的知识流动（溢出）。

从知识的角度看开源

数字公共产品的本质是：以数字形式凝聚的人类**知识**，这些知识可以以多种方式，节约人类的时间。

- 代码是一种知识。
- **Code Review**的过程是知识生产与传播的过程。
- 高质量的文档是一种知识。
- 社区内质量优秀的问答，也是一种知识生产与传播的过程。

在开源社区，知识的生产与传播过程，是一个整体。

- 知识的吸收者，同时在参与知识的生产过程。
- 使用开源软件并反馈的过程，同时是知识传播与促进知识生产的过程。
- 开源社区的交流氛围，会对于人类大脑里的知识，产生“虹吸”效应，促进知识的数字化。

开源生态的本质

- 开源生态的本质，就是由各种各样授权协议/规则所界定的、由全部数字公共产品组成的数字知识公地。所有参与这个开源生态的主体，基于共同遵守的规则，形成了开放利益共同体。
- 数字知识公地的本质：与公地悲剧中所说的物理世界的公地不同，只要数字世界依然存在，只要复制成本始终接近于0，数字世界的公地就不会减少，只会越来越繁荣。

从“知识所有权”到“知识共享”

存在知识产权保护的世界

知识经济
交易成本 >0

存在知识产权保护的世界

知识经济
交易成本 >0

开源世界

数字知识公地
知识的流动成本为零

开放式协作、知识共享与创新

开放式协作的主要特征

- 开放参与，来去自由
- 有明确的规则，但是也允许开放讨论，可以不断改进
 - 最基础的规则，是通过License规范的
- 基于数字公共服务，使协作过程，开放透明
- 独特的荣誉与激励机制
- 不断酝酿的创新，且与社区活跃度高度正相关

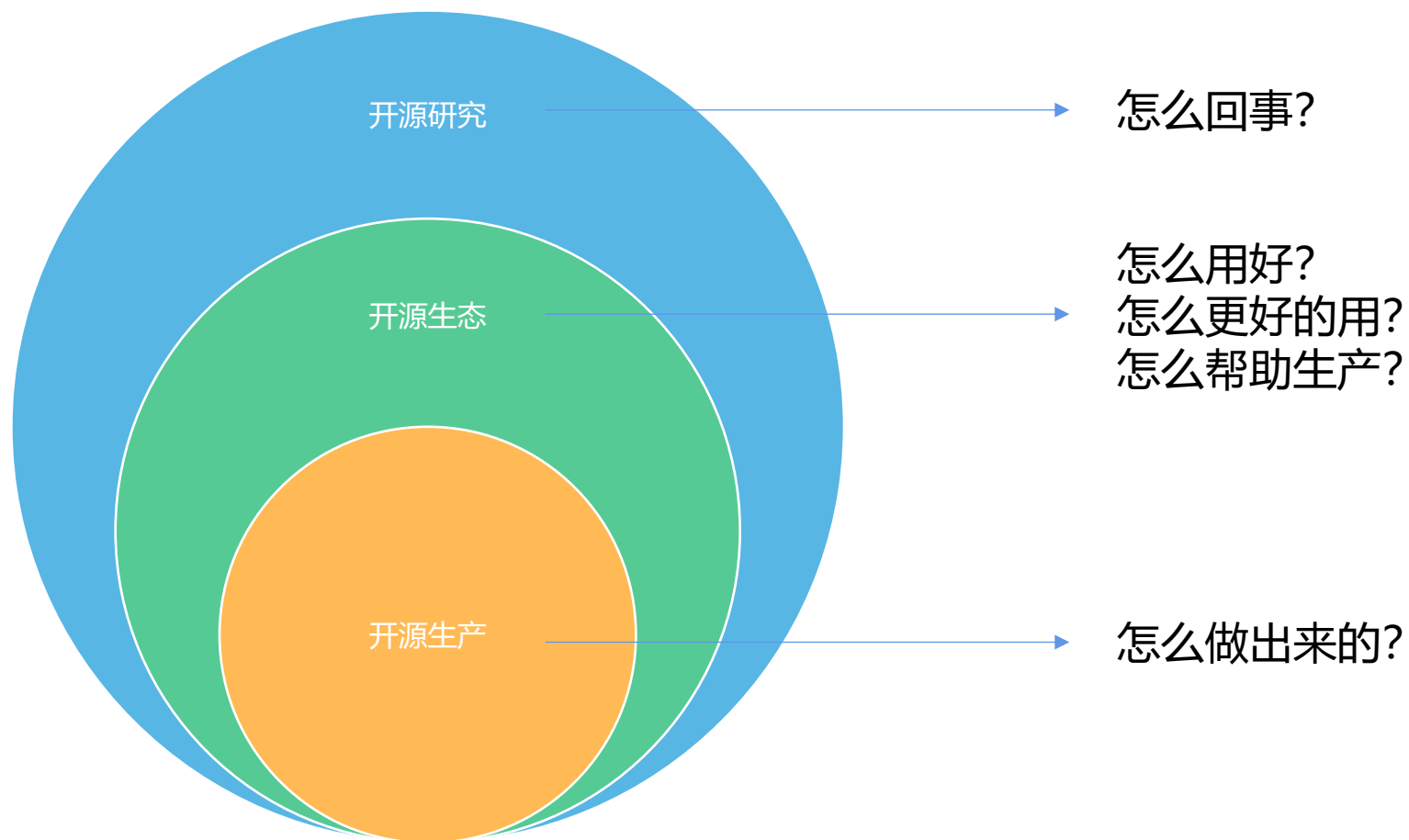
知识共享与创新

- 经过上述的分析，我们可以看出，在开源社区的协作方式中，形成了一种“知识共享”的生产与传播氛围
- 交易成本为0，非物质激励（礼物文化），鼓励分享/共享
- 这样的氛围与机制，极大的促进了知识的创新

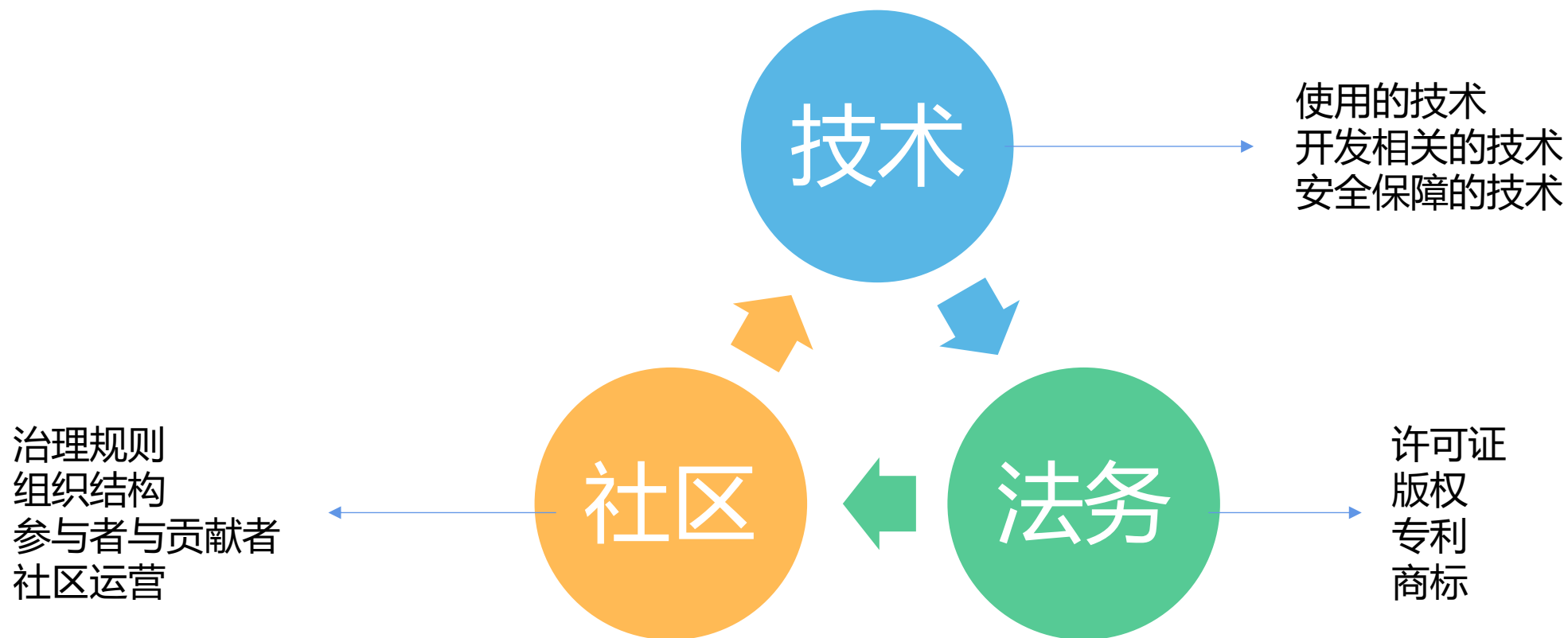
开源世界的知识图谱

如何更加全面的理解开源

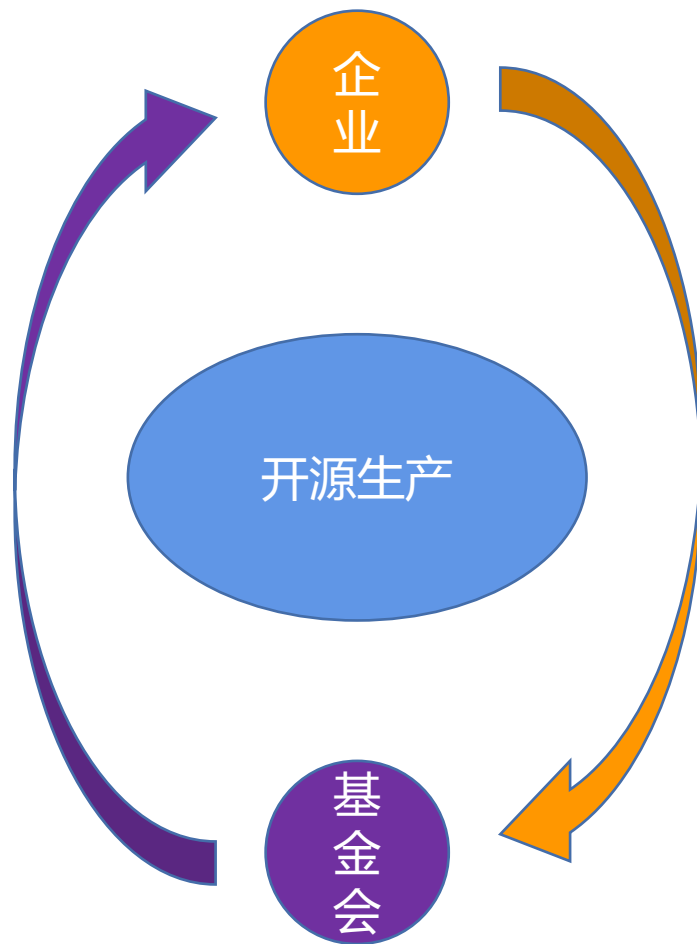
开源的内涵与外延



开源的生产过程

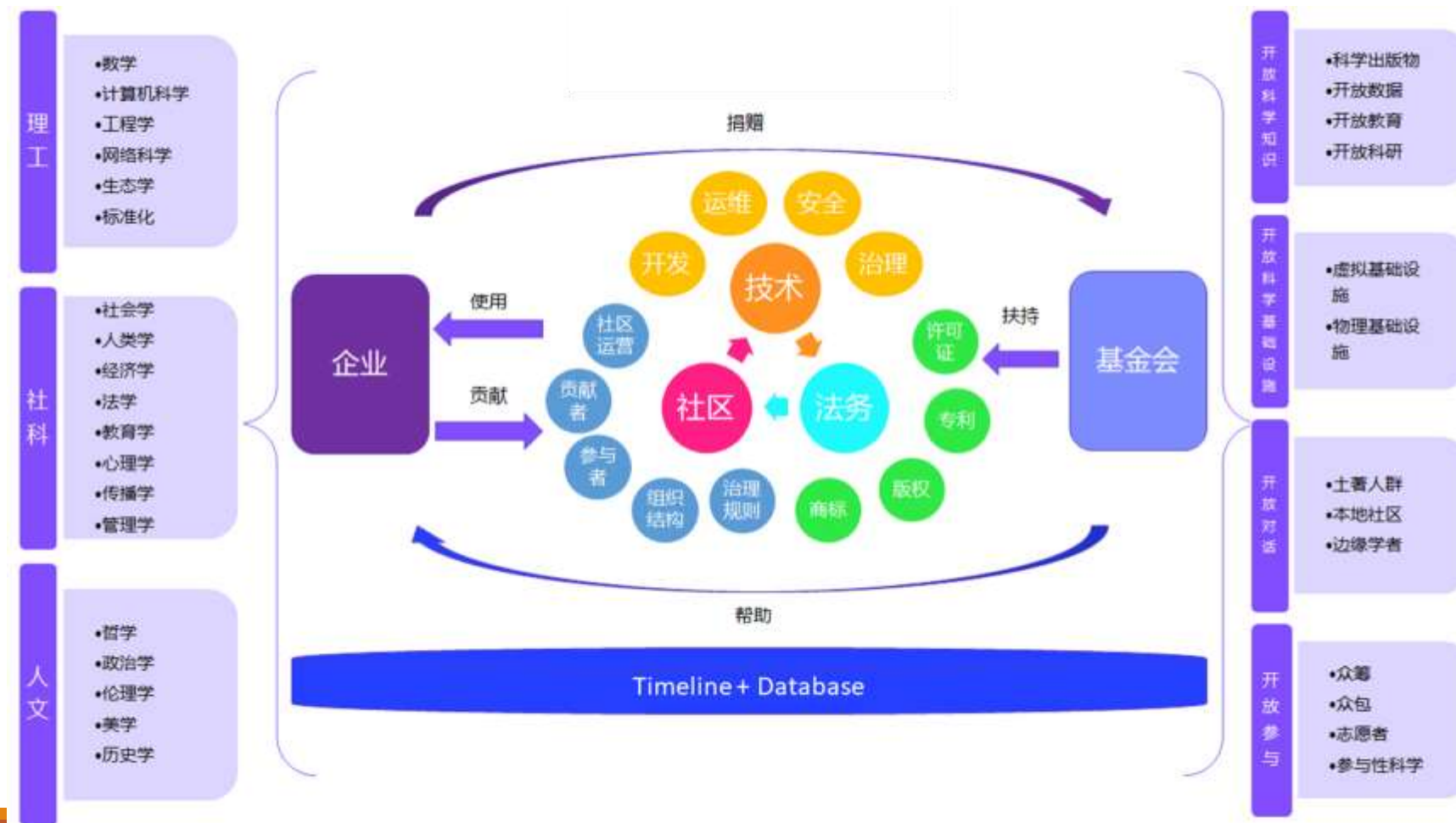


开源的生态环境

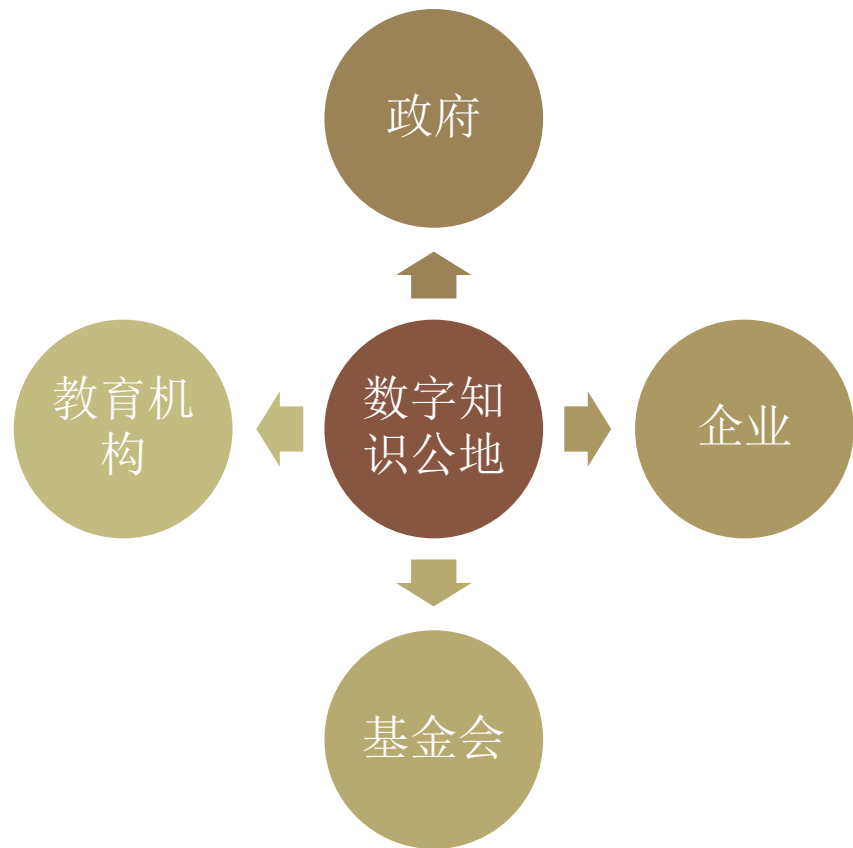


- 安全
- 有序
- 繁荣
- 有利

开源世界全景图



从知识公地的角度看开源世界：如何构筑健康的公益循环？



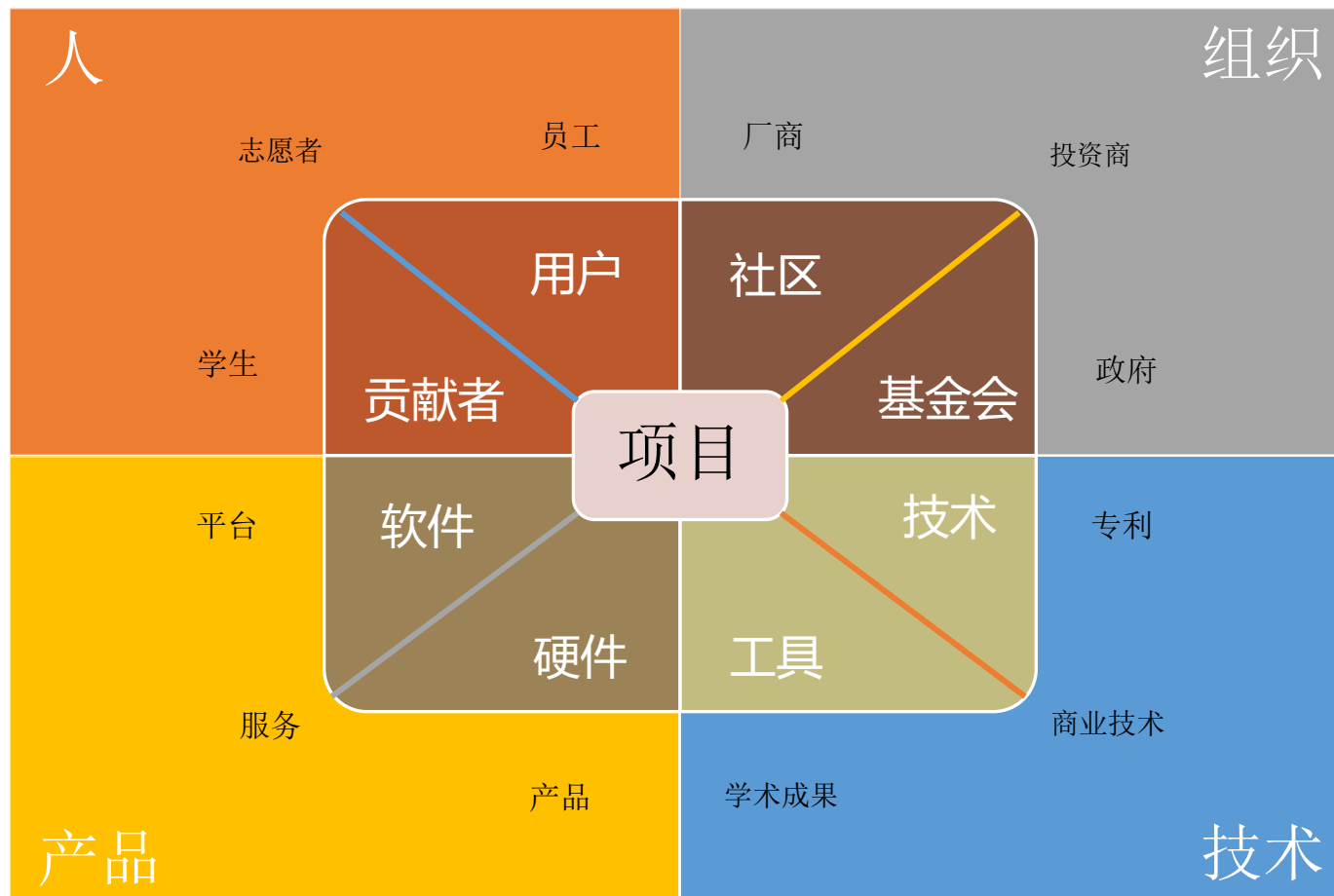
开源激励模型

- 外部激励
 - 利益
 - 商业、经济、管理
 - 名誉
 - 文化
- 内在激励
 - 乐趣
 - 文化
 - 能力
 - 教育

开源保障模型

- 不同层面的保障
 - 全球化层面
 - 国家层面
 - 组织层面
 - 企业层面
- 规则
 - 经济规则
 - 法律
 - 社区规则
- 潜规则
 - 社会风俗
 - 社区文化

为开源世界建模



当今开源面临的挑战

事情总不会那么一帆风顺

林纳斯眼球定律逐步失效

该定律表示：“足够多的眼睛，就可让所有问题浮现”（given enough eyeballs, all bugs are shallow）更正式地来说：“只要有足够的单元测试员及共同开发者，所有问题都会在很短时间内被发现，而且能够很容易被解决”。软件审查的一种形式是将代码展示多开发者以达成共识。代码审查能让研究者和开发者工作更有效率，且比软件测试更高效。

- 随着开源软件的数量越来越多，平均到每个项目上的关注者数量，就越来越少了。
- 由于对于开源软件的关注度存在聚焦效应，只有少量的开源项目获得了大量的关注，而大量的开源项目，则只有使用者，没有关注者。
- 这使得开源软件在质量、安全性方面，都存在失守的风险。

另一方面，礼物文化成立原因，也在于关注度，为开源项目贡献的大多数人，事实上却无人关注，更无人崇拜。

- 礼物文化的效果消失了
- 这就产生一种“冰山现象”，海平面以下的项目，海平面以下的开源贡献者，都“不再重要”。



商业开源与社区开源的激励机制产生矛盾

社区开源的动机是：Just for fun

商业开源的动机是：合理的商业回报

当一个商业开源企业，要运营一个开源社区时，不同动机的人会在一起为一个项目贡献，这就会变得很麻烦。

为何现在的很多开源项目，尤其是企业主导的开源项目，纷纷修改授权协议？

- 传统的礼物文化有点走不下去了
- 互联网的放大效应，导致头部与尾部的贡献者，收益差距越来越大
- 云计算平台的崛起，正在侵蚀开源创作者的利益

“后开放”授权协议，能不能走通？

- 开源定义（OSD）的撰写人，著名的 Bruce Perens 发布的一套新的授权协议
- 打算组织成立某种“开源开发者工会”，共同维护自己的权益

全球化共识快要失效了

全球化共识的假设：我们大家共建一个“公地”，大家都能从中受益。

- 美国人可能的假设是：“我肯定是受益最多的那个”
- 现在他们觉得不划算，这个事情他们就不想继续玩了。
- 国家利益的优先：美国等国家从自身利益出发，逐渐调整政策，限制某些技术的对外共享，尤其是在高科技和信息技术领域，导致全球合作的环境开始封闭。

知识封锁现象的出现

- **开源的封闭化利用：**开源最初是以自由共享和共同开发为核心理念，但随着全球竞争加剧，一些公司和国家开始在开源项目上施加限制，甚至对开源代码进行再封闭化，将成果私有化。这种做法违背了开源精神，形成了“知识封锁”现象，导致原本开放的资源重新被控制。
- **技术壁垒的加深：**在人工智能、5G、半导体等高科技领域，由于知识和技术的封闭化，国际间的技术流动和创新协作面临障碍。这种封闭趋势一方面加剧了技术壁垒，另一方面阻碍了其他国家和个人从开源项目中获取知识和技术。

知识封锁的影响

- **创新的集中化与不平等：**知识闭锁会导致创新资源集中于少数国家或企业，限制了全球范围内的创新活力。缺乏开放的知识共享机制，很多国家和个人将无法获得技术资源，导致创新资源分配不平等。
- **技术的武器化：**当知识和技术被封闭化后，可能被当作政治或经济手段进行操控，加剧了全球竞争与冲突。技术和知识的“武器化”意味着封锁技术的国家掌握了巨大的话语权和控制力，削弱了开放合作的可能性。



结论

开源的价值与启示



对知识生产与传播的启示

开源如何重塑知识的创造与流通模式？

- 知识的民主化：开源降低知识生产和获取的门槛。
- 创新模式的变革：基于开放的知识生产模式带来持续创新。

从上游到下游的
知识流动

知识的生产与传
播在一个平等的
社区内发生

针对知识的两种不同的观点

对社会发展的价值

开源对技术进步、教育普及与全球协作了哪些贡献？

- **技术进步：** 开源推动创新，提高技术普及度。
- **教育普及：** 开源资源在教育中的应用推动知识公平。
- **知识透明度：** 开源软件中蕴含的知识，都可以知其然、更知其所以然。
- **开放式协作：** 这是一套不仅可以用来编写软件的协作方法论。

未来展望

我们的行动倡议：

- 推而广之（**Advocate widely**）
 - 我们应该更加努力的向大众传播开源的理念与精神，让更多的人接受开源的理念，成为开源的同道中人。我们还应该在开源软件、开源硬件之外的领域，推广开源的实践——不仅开放源代码，而且开放数据，开放一切可以帮助这个世界变得更好的知识与经验，让更多的行业、更多的群体都接纳开源，成为开放式协作的受益者。
- 互帮互助（**Help each other**）
 - 我们应该帮助更多的开源项目，茁壮成长，臻于至善；帮助更多的开源社区，团结互助，共同协作。我们应该防止开源的含义被滥用或曲解。我们应该阻止割裂，反对人为设置的障碍，反对任何附加歧视条款的“伪开源”，确保开源始终是一项惠及全球的事业。
- 立即行动（**Just do it**）
 - 每一个人都可以参与开源，而不是只有大咖才能做到。我们可以从纠正拼写做起，翻译或撰写文档，为代码除错，提交代码，审核代码，志愿支持开源活动，甚至布道演讲，吸引更多的朋友加入。

面对一个急速变幻、急速下坠、甚至可能变得撕裂的世界，我们更应该推广开源精神，凝聚众人向善之力，汇聚众人向善之智，开放协作、互惠共赢，为这个世界变得更好做出贡献！

——《开源人宣言》

<https://github.com/kaiyuanshe/Open-Source-Fans-Manifesto>

谢谢