

第二章 开源协作的规则与法律框架

以社区文化与法律工具双轮驱动的协作生态系统

无形的文化与规范

开源文化的四大核心理念

开 开放 (Openness)

"只要有足够多的眼球，所有bug都将无处遁形" —— Linus定律
源代码对全球开放，最大化集体智慧，降低知识获取门槛

协 协作 (Collaboration)

"社区胜于代码" —— Apache基金会理念
跨越地理、时区与组织边界的分布式协作模式

共 共享 (Sharing)

知识、代码与经验的无障碍流动
避免"重复造轮子"，通过网络效应放大价值

贡 贡献 (Contribution)

循环的燃料，保持项目活力
多种形式：代码、文档、翻译、支持、测试等

协作模型的演进

大教堂模式

- 由少数精英构建
- 封闭式开发环境
- 精密的架构蓝图
- 源代码两次发布间保密
- 层级决策结构
- 典型案例：早期GNU项目

市集模式

- 大规模协作开发
- 完全公开的过程
- "发布早期，频繁发布"
- 用户即潜在开发者
- 快速迭代与反馈
- 典型案例：Linux内核

社区礼仪与行为准则

社区遵循精英治理(Meritocracy)原则，个体的声誉和影响力通过贡献的质量和持续性建立。

行为准则(Code of Conduct)将社区礼仪正式化，定义可接受与不可接受的行为，提供冲突处理机制，创造包容的环境。

典型社区角色包括：维护者(Maintainer)、贡献者(Contributor)和用户(User)，形成流动的结构。

有形的法律工具

出站许可(Outbound License)

出站许可定义了项目如何向外部世界授权使用其代码，回答了"用户可以用我的代码做什么？"这个问题。

互惠型(Copyleft)

GPL, AGPL, LGPL

要求衍生作品必须以相同许可证发布
确保代码自由传递，防止私有化
构建共享公地，强制所有改进回馈社区
形成紧密耦合、互惠互利的生态

宽松型(Permissive)

MIT, Apache 2.0, BSD

允许衍生作品以任何形式发布，包括闭源
仅要求保留版权声明
最大化代码采用率和行业影响力
形成"中心-辐射"式生态系统

入站贡献(Inbound Contributions)

入站贡献管理是项目如何合法、安全地接收外部贡献者的代码，确保项目知识产权链条的完整性。

CLA

贡献者许可协议

贡献者与项目间签署的正式法律合同
授予项目广泛的版权和专利许可
高法律确定性，简化再许可门槛高，但企业友好

DCO

开发者原创声明

通过Git提交中的Signed-off-by声明
贡献者证明有权提交代码
极低贡献门槛，对开发者友好
社区友好，但法律保护较弱

知识产权三要素

版 版权(Copyright)

一切的基础，许可证的核心
保护代码作为创作作品的表达形式

专 专利(Patent)

保护软件功能的实现方法
现代许可证(如Apache 2.0)含
专利授权与报复条款

商 商标(Trademark)

保护项目名称、标识与品牌
开源许可几乎从不授予商标权

开源协作的整体法律框架

开源项目的法律框架，是由出站许可与入站贡献机制共同构成的知识产权治理体系。
核心原则：项目从贡献者那里获得的入站权利，必须足以支撑它向用户做出的出站许可承诺。

贡献者

→

项目

→

用户

入站策略与出站许可的战略协同

项目类型	出站许可	入站贡献	战略逻辑
企业"开放核心"项目	宽松型(Apache 2.0)	CLA	保留将社区贡献整合到专有企业版的权利
社区驱动的基础设施项目	互惠型(GPL)或宽松型	DCO	最大化社区参与，降低贡献门槛
中立基金会"大帐篷"项目	宽松型(Apache 2.0)	CLA	为竞争公司创造安全的法律"避风港"

从代码到创作、数据与智能的新一代开源许可证

开源许可证的原则和法律模式已超越纯软件领域，扩展到硬件、文化创意、数据和人工智能等新兴领域。
每个领域都催生了适应其独特法律和技术挑战的新型许可证。

硬 开源硬件许可 (OSH)

开源硬件是指物理对象的设计规范，其许可方式允许任何人学习、修改、制造和分发这些设计及其衍生产品。

与软件许可的根本区别

开源硬件许可必须同时处理**版权**和**专利**两种知识产权：

- 版权保护设计文档的表达形式
- 专利保护功能性发明本身，是OSH许可的核心

CERN开放硬件许可(OHL)系列

许可类型	特点	适用场景
CERN-OHL-S (强互惠型)	要求整个使用该组件的系统都必须开源	确保整个硬件生态开放
CERN-OHL-W (弱互惠型)	仅要求对组件本身的修改开源，允许与闭源系统集成	平衡开放与商业化，适合模块化硬件
CERN-OHL-P (宽松型)	只要求保留版权声明，无互惠要求	最大化硬件设计的行业采用

"开源硬件的执行更依赖社区规范与创客精神，而非单纯的法律强制。"

创 知识共享许可 (Creative Commons)

CC许可为创造性作品提供了一套模块化的"保留部分权利"框架，促进了全球范围内的知识、文化和教育资源的共享。

四大构建模块

BY (署名)

SA (相同方式共享)

NC (非商业性使用)

ND (禁止演绎)

六种标准许可

- CC BY** - 最宽松，仅要求署名
- CC BY-SA** - 维基百科使用，要求衍生作品以相同许可发布
- CC BY-NC** - 允许修改，但限于非商业用途
- CC BY-NC-SA** - 非商业用途，且衍生作品需保持相同许可
- CC BY-ND** - 允许任何目的使用，但禁止修改
- CC BY-NC-ND** - 最严格，仅允许非商业性的完整分发

标准化的三层结构

- 人类可读的"许可简本"(Deed)
- 律师可读的"法律文本"(Legal Code)
- 机器可读的元数据

"CC许可催化了用户生成内容和开放教育资源的爆发式增长，实现了版权管理的民主化。"

数 开放数据许可

开放数据许可解决了数据和数据库共享的独特法律挑战，处理版权和数据库特殊权利问题。

数据的法律独特性

- 单个事实不受版权保护
- 数据库结构可能受版权保护
- 欧盟等地区存在"数据库特殊权利"，保护数据收集的投资

开放数据共享(ODC)许可套件

- ODbL** - 开放数据库许可，要求衍生数据库也开源(类似GPL)
- ODC-By** - 仅要求署名的宽松许可(类似MIT)
- PDDL** - 公共领域贡献，放弃所有权利

关键区别：数据库 vs. 数据库内容

ODC许可适用于数据库本身("容器")，而内部的单个内容("内容物")可能受其他许可制约，这种潜在的"许可陷阱"要求使用者必须进行双重验证。

"ODC许可为全球数据共享提供了稳健的法律基础，但使用者需警惕'容器'与'内容'的许可差异。"

智 开源人工智能(AI)许可

为AI系统许可是复杂的挑战，因其由源代码、训练数据和模型参数三种不同性质的资产构成。

AI资产的三元结构

- 源代码**: 受版权保护，适用传统开源许可。
- 训练数据**: 复杂的集合，可能包含第三方权利。
- 模型参数**: 核心产物，其版权地位存在激烈争议。

新兴许可模式与争议

- 负责任AI(RAIL)许可**: 增加基于使用场景的道德和安全限制。
- Meta Llama 许可争议**: 因包含商业使用限制和可接受使用政策(AUP)而被指责为"开放洗白"(open-washing)，违反了开源定义(OSD)。
- OSI的前进之路**: 正在制定严格的《开源AI定义》，要求提供复现所需的所有组件。

"AI领域的许可问题揭示了传统开源'绝对自由'与新兴'负责任使用'思潮间的深刻理念分裂。"