

团体标准

T/CESA 1270.5—2023

信息技术 开源治理 第5部分：开源贡献者 评估模型

Information technology—Open source governance—Part 5: Open source contributor
assessment model

2023-09-28 发布

2023-09-28 实施



版权保护文件

版权所有归属于该标准的发布机构，除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

目 次

前 言	V
引 言	VI
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 贡献者评估模型	1
6 技术级贡献	2
6.1 概述	2
6.2 新开议题	3
6.3 议题评论	3
6.4 新开合并请求	3
6.5 代码评审	4
6.6 代码合并	4
7 治理级贡献	4
7.1 概述	4
7.2 角色类型	5
7.3 社区传播	5
8 贡献者评估规则	5
8.1 概述	5
8.2 贡献范围	5
8.3 评估规则	6
附 录 A （资料性） 开源术语中英文对照表	7
附 录 B （资料性） 贡献度统计评估计算示例	8
附 录 C （资料性） 贡献度网络评估计算示例	9

前 言

本文件按照GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是T/CESA 1270《信息技术 开源治理》的第5部分。T/CESA 1270已经发布了以下部分：

- 第1部分：总体框架；
- 第2部分：企业治理评估模型；
- 第3部分：社区治理框架；
- 第4部分：项目评估模型；
- 第5部分：开源贡献者评估模型。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由中国电子技术标准院研究院提出。

本文件由中国电子技术标准化研究院、中国电子工业标准化技术协会归口。

本文件起草单位：中国人民解放军国防科技大学、中国电子技术标准化研究院、开放原子开源基金会、中兴通讯股份有限公司、浪潮电子信息产业股份有限公司、华东师范大学、阿里云计算有限公司、苏州棱镜七彩信息科技有限公司、浙江九州未来信息科技有限公司、上海计算机软件技术开发中心、蚂蚁集团集团股份有限公司、北京百度网讯科技有限公司、中移（苏州）软件技术有限公司、普元信息技术股份有限公司、北京华胜天成科技股份有限公司、中国软件评测中心（工业和信息化部软件与集成电路促进中心）、浪潮云信息技术股份公司、东软集团股份有限公司、中移系统集成有限公司、深圳市金蝶天燕云计算股份有限公司、湖南先进技术研究院。

本文件主要起草人：余跃、杨丽蕴、李成双、于秀明、李博、王涛、张洋、李响、张百林、王伟、龚斌、张迅晖、但吉兵、章津楠、王超、郭智慧、马红伟、钱岭、孟庆余、梁钢、袁薇、黄先芝、赵赫、李智琪、于昕、杨佳丽、邝敏越、杨程、周济、吴涛、曾晋、梁大功、彭晋、边思康、王旭、杨静、孙福洲、鞠东颖、黄浩东、王媛媛、魏弋钧、田晓利、水超、许绍楠、蒲天阳、郭雪雯、郭涛。

引 言

为规范开源领域治理活动，提升开源治理能力，系统性地指导开源各组织开展开源治理策略与政策编制，结合开源治理实际情况，制定开源治理标准。T/CESA 1270旨在为开源治理的事项与活动提供依据，拟由5个部分构成。

- 第1部分：总体框架。目的在于为不同开源参与组织确立完善、规范的开源治理框架和组成要素。
- 第2部分：企业治理评估模型。目的在于规定企业在自身开源治理过程中具备的方法、流程和能力。
- 第3部分：社区治理框架。目的在于描述了开源社区治理与运营总体框架的利益相关方、治理原则、框架与组成要素、成熟度模型、成熟度评估和数据评价方法。
- 第4部分：项目评估模型。目的在于确立了建立开源项目多维度指标模型，系统性评价开源项目发展情况。
- 第5部分：开源贡献者评估模型。目的在于确立了开源贡献者参与开源技术贡献和治理贡献的评估模型。

信息技术 开源治理 第5部分：开源贡献者评估模型

1 范围

本文件确立了开源贡献者评价模型，规定了技术级贡献和治理级贡献两个评估域及评估规则。本文件适用于开源参与主体实施开发、运营等人员的贡献度评价、能力评估和团队管理。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

T/CESA 1269-2023 信息技术 开源 术语与综述

3 术语和定义

T/CESA 1269-2023 界定的以及下列术语和定义适用于本文件。

3.1

技术级贡献 technical level contribution

贡献者在开源社区中进行的所有与代码、议题等直接相关的贡献行为。

3.2

治理级贡献 governance level contribution

贡献者在开源社区中进行的所有与运营治理任务相关的贡献行为。

4 缩略语

下列缩略语适用于本文件。

LOC: 代码行数 (Lines of Code)

OSC: 开源社区 (Open Source Community)

OI: 新开议题 (Open Issues)

IC: 议题评论 (Issue Comments)

PR: 合并请求 (Pull Request)

OP: 新开合并请求 (Open Pull Request)

PRR: 合并请求评审 (Pull Request Reviews)

MP: 合并请求被接纳 (Merge Pull Request)

5 贡献者评估模型

贡献者评估模型由技术贡献和治理贡献两个贡献领域构成。贡献者评估模型的评估域、评估子域和评估项组成见图1。

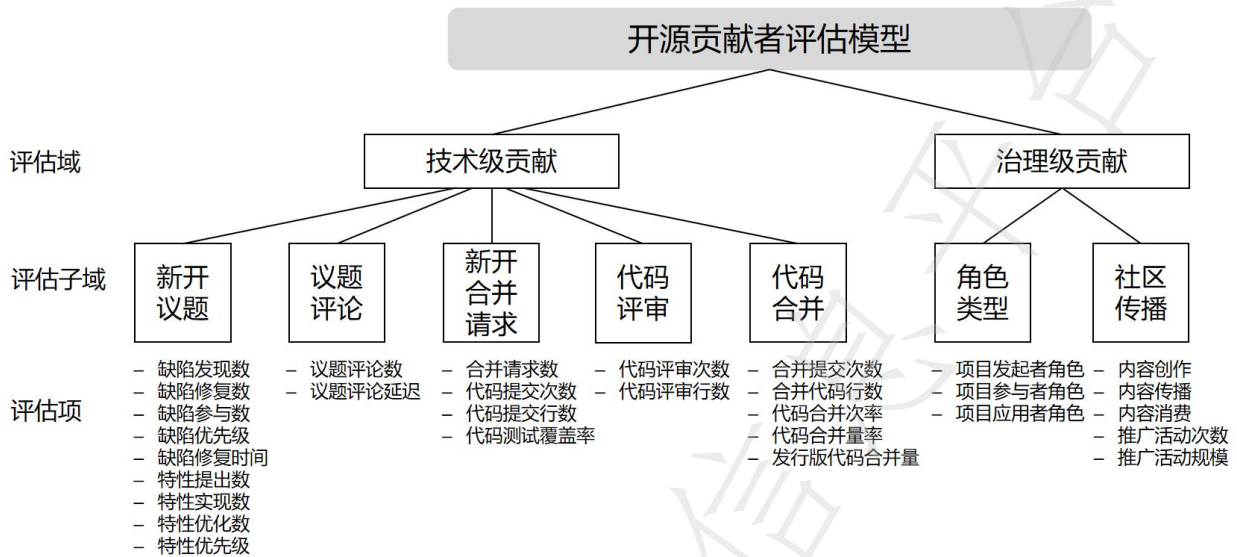


图 1 开源贡献者评估模型

6 技术级贡献

6.1 概述

技术级贡献是开源贡献者对开源项目的最基本贡献类型，泛指开发者在开源项目中进行的所有与代码、议题直接相关的技术贡献行为。技术级贡献包括新开议题、议题评论、新开合并请求、代码评审和代码合并五个评估子域及22个具体评估项，技术级贡献评估内容见表1。相关术语中英文对照见附录A。

表 1 技术级贡献评估内容

评估子域	评估项
新开议题	缺陷发现数
	缺陷修复数
	缺陷参与数
	缺陷优先级
	缺陷修复时间
	特性提出数
	特性实现数
	特性优化数
	特性优先级
议题评论	议题评论数
	议题评论延迟

表1 技术级贡献评估内容（续）

评估子域	评估项
新开合并请求	合并请求数
	代码提交次数
	代码提交行数
	代码测试覆盖率
代码评审	代码评审次数
	代码评审行数
代码合并	合并提交次数
	合并代码行数
	代码合并次数
	代码合并量率
	发行版代码合并量

6.2 新开议题

新开议题包括：

- a) 缺陷发现数：反映开发者在缺陷发现方面的贡献，统计向开源社区提交且被确认有效的缺陷报告数量；
- b) 缺陷修复数：反映开发者在解决缺陷的能力与贡献，统计解决开源社区中有效缺陷的数量；
- c) 缺陷参与数：反映开发者在协同解决缺陷过程中的贡献，统计参与讨论或解决的缺陷数量；
- d) 缺陷优先级：反映开发者解决不同优先级缺陷的贡献，根据社区缺陷优先级的定义，从高到低设置不同分值，并根据所解决缺陷优先级统计总体得分；
- e) 缺陷修复时间：反映开发者解决不同项目缺陷的效率，即从接受缺陷指派到解决缺陷并提交代码的时间。注意该项只考虑被社区认定的、可跟踪的缺陷；
- f) 特性提出数：反映开发者在开源项目特性需求方面的贡献，统计项开源社区提交且被认可的特性需求请求数量；
- g) 特性实现数：反映开发者实现新功能、新特性的能力和贡献，统计开发完成的特性需求次数；
- h) 特性优化数：反映开发者在特性协同实现过程中的贡献，统计参与优化或实现的特征需求数量；
- i) 特性优先级：反映开发者实现不同优先级特性需求的贡献，根据社区特性需求优先级的定义，从高到低设置不同分值，并根据所实现优先级统计总体得分。

6.3 议题评论

议题评论包括：

- a) 议题评论数：反映开发者参与议题的讨论与交流情况，即统计参与的议题评论总数；
- b) 议题评论延迟：反映开发者参与议题讨论与交流的响应速度，即统计平均评论回复延迟。

6.4 新开合并请求

新开合并请求包括：

- a) 合并请求数：反映开发者在合并请求提交方面的贡献，统计向开源社区提交且被确认有效的合并请求数量；
- b) 代码提交次数：反映开发者向开源项目主动贡献的代码变更数量，即统计提交的代码变更总数；
- c) 代码提交行数：反映开发者向开源项目主动贡献的代码量情况，即统计提交的代码行数总和；
- d) 代码测试覆盖率：从代码自动化用例（或测试用例）的覆盖率角度反映开发者向开源项目贡献优质代码的能力水平，即计算某个周期内提交代码的平均自动化用例覆盖率。

6.5 代码评审

代码评审包括：

- a) 代码评审次数：反映开发者参与议题讨论与交流体关注情况，即统计参与代码评审的总次数；
- b) 代码评审行数：反映开发者对开源项目中其它贡献者的具体关注程度，即统计评审的代码总行数。

6.6 代码合并

代码合并包括：

- a) 合并提交次数：反映开发者被开源项目接受的代码变更数量，即统计代码仓库中被合并的代码变更总数；
- b) 合并代码行数：反映开发者被开源项目接受的优质代码总量，即统计代码仓库中被合并的代码行数总和；
- c) 代码合并次率：从代码变更的角度反映开发者向开源项目贡献代码的能力水平，即计算代码合并次数与代码提交次数的比值；
- d) 代码合并量率：从代码量的角度反映开发者向开源项目贡献代码的能力水平，即计算代码合并行数与代码提交行数的比值；
- e) 发行版代码合并量：从进入发行版代码量的角度反映开发者向开源项目贡献代码的能力水平，即计算代码最后进入社区发行版中的代码行数。

7 治理级贡献

7.1 概述

治理级贡献是开源贡献者对开源项目的重要贡献类型，泛指贡献者在开源项目中进行的所有与运营治理任务相关的贡献行为。治理级贡献包括角色类型和社区传播两个评估子域，其中角色类型指贡献者在特定开源项目中的参与角色行为，社区传播泛指在开源社区中扩大项目传播范围、提高开源项目影响力等相关的贡献行为。治理级贡献评估内容见表2。

表2 治理级贡献评估内容

评估子域	评估项
角色类型	项目发起者角色
	项目参与者角色
	项目应用者角色

表2 治理级贡献评估内容（续）

评估子域	评估项
社区传播	内容创作
	内容传播
	内容消费
	推广活动次数
	推广活动规模

7.2 角色类型

角色类型包括：

- a) 项目发起者角色：反映开发者是项目发起者的角色进行贡献；
- b) 项目参与者角色：反映开发者是以参与项目开发的角色进行贡献；
- c) 项目应用者角色：反映开发者主要作为使用者的角色进行贡献。

7.3 社区传播

社区传播包括：

- a) 内容创作：反映开发者为开源项目贡献内容的活跃度，采用近一年的内容发布数量表示，如文章、发帖、视频等；
- b) 内容传播：反映开发者对开源项目传播的贡献程度，采用内容传播次数表示，如转发、分享链接等；
- c) 内容消费：反映开发者对开源项目内容的关注度，采用内容消费次数表示，如浏览、阅读、点赞等；
- d) 活动次数：开发者作为主办方或主要参与方在线/线下举办或参加的开源技术交流、开源项目推广等活动的次数；
- e) 活动规模：开发者作为主办方或主要参与方在线/线下举办或参加的开源技术交流、开源项目推广等活动的规模，采用活动受众的总人数表示。

8 贡献者评估规则

8.1 概述

本章给出具体的基于代码托管平台的开源贡献者评估规则，作为贡献者评估模型的计算方式。代码托管平台是当前主要的开源贡献发生场所，且基于Git开发模式的代码托管平台已经成为全球主流，本章的评估规则均是基于Git类代码托管平台，围绕贡献度统计规则和贡献度网络评估规则两个维度进行规范。

8.2 贡献范围

根据本文件所界定的技术级贡献，结合代码托管平台的实际功能，将参与过如下五类贡献的人员认定为贡献者，包括：

- a) OI：新开一个议题，用于开展日常的开发事务；
- b) IC：在某一个议题中进行日常讨论与交流；

- c) OP: 新发起一个代码合并到仓库中的请求;
- d) PRR: 代码合并过程中的评审与讨论;
- e) MP: 代码最终被合并到代码仓库中。

8.3 评估规则

8.3.1 贡献度统计评估规则

贡献度统计评估规则及计算方法: 本规则将8.2中五类事项在某一时间段中发生的数量, 通过加权求和的方式进行计算, 进而对一个开发者参与开源的贡献程度进行量化。即:

$$S = a*OI + b*IC + c*OP + d*PRR + e*MP$$

式中:

S ——统计评价模型的社区活动计算结果;

OI ——新开议题数量;

IC ——议题评论数量;

OP ——新开合并请求数量;

PRR ——合并请求评审数量;

MP ——完成合并请求数量;

a 、 b 、 c 、 d 、 e ——每类事项的价值权重, 一般由用户通过专家共识的方式完成。

具体计算示例见附录B。

8.3.2 贡献度网络评估规则

贡献度网络评估规则及计算方法: 本规则利用8.2中五类贡献的全部信息, 在某一时间段内, 通构建开发者协作关系网络, 得到一个全局的包含“开发者与项目”、“开发者与开发者”、“项目与项目”之间的贡献关系网络。在贡献关系网络上通过大规模图算法, 就可以得到在该时间段, 所有开发者参与到所有开源项目贡献中的贡献度。具体示例可见附录C。

附录 A
(资料性附录)
开源术语中英文对照表

相关开源术语中英文对照见表A.1。

表 B.1 开源术语中英文对照表

序号	术语 (中文)	术语 (英文)
1	代码仓库	repository
2	开发者	developer
3	开源项目	open source project
4	代码提交	commit
5	代码行数	lines of code
6	合并请求	pull-request
7	缺陷	bug、defect
8	议题	issue
9	特性	feature
10	帖子	post
11	评论	comment

附 录 B
(资料性附录)
贡献度统计评估计算示例

具体场景如高校师生参与开源项目的贡献度评估，可将议题评论每个计1分、新开议题每个计2分、新开合并请求每个计3分、提交合并请求评审每个计4分、合并请求被接纳每个计5分，可得详细开发者贡献度统计评估指标体系，见表B.1，其中二级指标的权重，可根据具体场景自主调节。

表 B.1 开发者贡献度统计评估指标体系

一级指标	二级指标	权重
贡献度统计评估	OI	5%
	IC	10%
	OP	20%
	PRR	25%
	MP	40%

附 录 C
(资料性附录)
贡献度网络评估计算示例

开源贡献协作网络构建基本逻辑为：在全域贡献协作网络中，以项目与开发者为节点，以贡献关系为边可构建一个贡献协作网络，其含义为，以项目作为开发者的协作单元，即在同一个项目上贡献过则为一种协作。在项目内部，使用Issue和PR作为基本的协作单元，在同一个Issue或PR进行的讨论就被看成是一种协作。则可得到一个项目内的协作网络图，构建一个开源项目内的精细化协作贡献网络。见图 C.1。基于此网络，可以利用一些图分析的算法来计算每个项目、以及每个开发者的协作贡献度，例如PageRank算法。

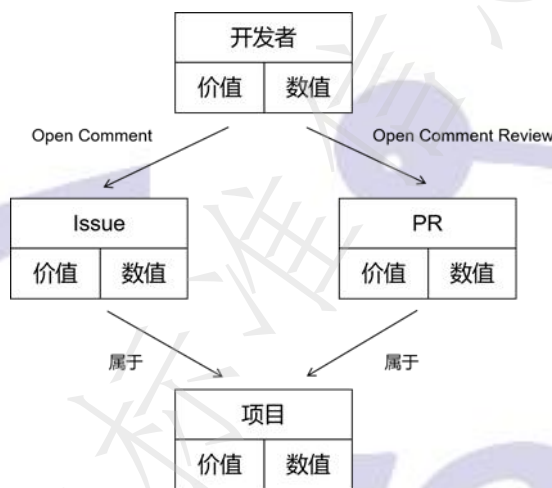


图 C.1 项目内部协作网络

参考文献

- [1] Kononenko O, Rose T, Baysal O, et al. Studying pull request merges: a case study of shopify's active merchant[C]//Proceedings of the 40th international conference on software engineering: software engineering in practice. 2018: 124-133.
- [2] Hemmati H. How effective are code coverage criteria?[C]//2015 IEEE International Conference on Software Quality, Reliability and Security. IEEE, 2015: 151-156.
- [3] McIntosh S, Kamei Y, Adams B, et al. An empirical study of the impact of modern code review practices on software quality[J]. Empirical Software Engineering, 2016, 21: 2146-2189.

