附录 B

模型开放条件与数字公共品研究参照

开放材料为共同改进提供条件，公共价值还需要具体使用与持续维护的证据。本附录承接第五章的模型开放观察与第六章的行动建议，说明怎样界定判断对象、组织版本证据，并区分报告分类、开放定义与数字公共品认可。附录 A 解释指标、样本和统计方法，本附录关注材料、权利与公共用途之间的连接。

B.1 三类判断与各自依据

阅读“开放模型”或“数字公共品”时，首先确认这句话依据哪套规则、针对哪个对象。模型家族名称、某个权重版本与包含服务和数据的完整解决方案，其范围并不相同。

判断层次	回答的问题	应采用的依据
报告中的开放观察	提供了哪些材料，源评估如何分类?	5.3 的 L/O 释义、源评估结果及版本证据
OSI 开源 AI 定义	是否具备使用、研究、修改与分享的自由及所需材料?	OSAID 1.0 的定义及适用条款
DPGA 正式认可	申报的数字解决方案是否通过标准评审?	DPG Standard、官方评审和有效 Registry 记录

5.3 分别呈现材料完整性 L 与许可开放性 O，高完整、高开放是报告的结构比较分组。它们与 OSI 定义、DPGA 认可没有自动换算关系。本报告不授予认证，也不设置总分、金银等级或自定义的公共品合格线。

本附录核对的官方参照截至 2026 年 10 月 22 日；案例沿用 2026 年 10 月 17 日的冻结记录。标准更新、案例更新与报告修订分别记录，不能用新的核对日期覆盖旧证据的时间。

B.2 固定版本，逐项查证开放材料

查证从一个具体发布单元开始：保留完整模型 ID、发布者、提交或版本号、材料入口与核验日期。上游模型、数据和运行依赖分别列明；同一模型家族的其他尺寸、微调版和后续版本不自动继承结论。

本报告用六类材料组织查证。表中是研究记录的内容要求，不能直接代替任何标准的完整评审清单。

材料类别	需要留下的记录
权重与配置	文件、配置、分词器及获取条件；说明对应版本和检查范围
推理实现	实现与依赖版本、运行入口、环境要求及已执行的任务
完整预训练代码	数据处理、训练配置与运行链；区分预训练、微调和推理代码
训练数据信息	来源、组成、获取方式及处理规则；说明公开数据、受限数据和仍未知的部分
许可与使用条件	分别记录代码、权重、数据及依赖的条款，保留具体文本与日期
维护与反馈	维护主体、问题入口、版本变更与迁移记录；入口存在不代表响应质量已验证

“已观察”表示查到材料或声明，具体完整程度仍需阅读详情；“待核实”表示证据不足；“不适用”需要说明理由。实际运行另记环境、任务、结果和失败原因。六类记录不加权求和，材料可见与独立验证通过分别呈现。

OSAID 1.0 将四项自由与修改所需的数据说明、完整训练及运行代码、参数和相应开放条款相联系。详细数据信息须足以支持有能力者构建实质等效系统，并按定义交代数据来源与获取方式；它不等于笼统要求重发全部原始数据。仅能下载权重，或仅看到 Apache-2.0 等许可名称，都不足以判定整个系统符合定义。

B.3 四个固定版本案例的阅读方法

第五章保留四个有目的选择的案例，用于展示证据怎样组织，不构成代表性样本。下表概括 2026 年 10 月 17 日的记录；完整提交号、逐字段来源和核验说明保留在版本档案中。

固定版本案例	本次记录支持怎样的判断
Qwen2.5-7B · d1497293	权重和推理入口可查，许可记录为 Apache-2.0；完整预训练运行链仍待核实，概要数据介绍不代替完整材料
OLMo-2-1124-7B · 7df9a825	权重、分阶段训练配置与具名数据入口形成较完整的材料链；报告未重跑完整训练，数据组件的许可仍需分别核对
Llama-3.1-8B · d04e592b	权重访问与运行入口可查，需结合该版本社区条款及使用条件；不能由文件可见推定获得不受限的使用权利
Gemma-2-9B · 33c19302	权重获取需接受条款，数据与处理方式有概要说明；完整预训练链待核实，适用条件按档案中的条款日期定位

提交号固定的是模型仓库，外部许可、数据集与维护页面可能另有版本。引用时应同时保留这些来源的日期，不能把模型仓库的一个提交号当作所有依赖均已冻结。

四个案例不加入 870 个报告样本或 884 条平台记录，也不替代源平台的等级。两套样本未完成身份对齐，差值不代表增长；884 条记录中的 22 条许可字段差异继续保留待复核状态。相关统计边界见附录 A，逐项查证入口见5.3 模型证据档案。

B.4 数字公共品标准与正式认可

数字公共品认可面对的是申报的数字解决方案。DPGA 的九项指标覆盖公共目标、开放条件、治理与伤害防范；下表概括查证方向，正式要求以DPG Standard及其适用类别说明为准。

标准指标	查证方向
1 SDG 相关性	对应的公共问题、发展目标与支持证据
2 开放许可	各类组成部分是否采用认可的开放许可或条款
3 清晰权属	谁拥有、管理并对相关资产负责
4 平台独立	必需依赖及其限制，是否具备开放替代方案
5 文档	他人理解、部署、使用和改进所需的说明
6 非个人数据提取	适用时能否以开放格式导入或导出非个人数据
7 隐私与适用法律	隐私安排、数据处理规则与法律遵循说明
8 标准与最佳实践	互操作、可访问性及相关技术实践
9 设计上的伤害防范	数据隐私与安全、内容治理及反骚扰机制

AI 系统还需核对专门的组成部分要求。DPGA 在 2025 年更新中要求开放训练、验证和测试数据，以及相应代码、模型架构和参数；其数据要求不能由 OSAID 的“详细数据信息”替代。模型与数据文档、风险及缓解说明也属于官方要求，参见AI 系统说明。

正式认可以官方评审及 DPG Registry 的有效记录为依据。当前规则规定认可有效期为一年，之后需重新评审；引用时注明条目、评价日期与查询日期。联盟成员资格与项目认可分别查证，不把机构身份转写为其全部产品的认可。

DPGA 的评审聚焦核心解决方案设计，不涵盖每一个本地部署。获得认可仍需在具体环境中验证适用性、维护能力与实际效果；本报告的公共价值观察也不能替代这一评审。

B.5 把开放条件连接到公共价值

第一章提出持续共同创造的价值方向，第六章将其落实为共同任务。对具体模型或解决方案，还要回答谁能使用、怎样参与、是否解决问题，以及谁承担后续维护。下载、榜单位次或材料数量可以帮助寻找对象，不能单独证明这些成效。

下表是本报告建议的研究与交付记录，不是新增的 DPGA 认证条款。项目应根据用途选取可验证的指标，避免为了统一评分而忽略使用环境。

公共价值问题	建议保留的证据
谁能够实际使用?	目标人群、语言和可访问性需求、部署成本、硬件条件及未覆盖群体
是否解决了具体问题?	任务与基线、评价范围、使用结果、已知局限，以及对结果的独立核对
谁能够继续改进?	贡献入口、文档与验证任务、问题处理过程、下游修改及回馈记录
如何保持长期可用?	维护责任、资源安排、版本迁移、依赖更新与退出或交接方案

同一方案在不同语言、地区和资源条件下，可能产生不同效果。评估应写明时间、人群、任务与比较方法；没有对照或足够证据时，将观察结果与因果解释分开。多语言支持、轻量部署与共同维护可作为发展任务，不能统一写成所有项目必须通过的官方门槛。

B.6 交付、更新与集中查证

一次可复核的交付应让后续团队知道依据是什么、哪些工作已经完成，以及下一步由谁处理。建议保留以下四组记录，并随版本变更持续维护。

记录组	最少保留的信息
对象与依据	模型或解决方案边界、发布者、完整版本、上游关系、适用规则及其版本
材料与判断	材料及许可来源、核验日期、观察状态、验证任务和结果、尚未解决的问题
用途与责任	服务场景、目标人群、已知限制、维护与反馈负责人、后续验证安排
变更与复核	修改内容与原因、影响的旧结论、复核者与日期；涉及认可时保留有效官方记录

材料、许可、依赖、标准或用途发生变化时，先确定哪些判断受到影响，再更新相应记录。新版本的发布不自动证明旧版本已经补齐；旧证据应可追溯，待核实项在取得证据后再变更状态。

全书的查证分工是：附录 A 解释指标、时间、样本与比较边界；附录 B 解释开放材料、正式认可与公共用途。沿着具体图号回到来源，再按实际任务开展验证，才能把报告观察转化为可持续的共同工作。

报告入口：附录 A、5.3 开放观察、第六章行动建议。完整案例记录见版本档案。

官方参照：OSI OSAID 1.0、DPG Standard、AI 系统要求、评审与续期规则。本附录为研究归纳，具体申请按官方现行要求办理。