

团 体 标 准

T/DGAG XXX—XXXX

基于人工智能的职业教育教学质量评价 数据共享规范

Standard for data sharing of quality evaluation system of vocational education
based on artificial intelligence

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

广东省数字政务协会 发 布

目 次

前言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 缩略语 2

5 总体架构 2

 5.1 总体架构 2

 5.2 共享层级与范式 3

6 共享要求与策略 3

 6.1 共享基本原则 3

 6.2 共享层级与范式 4

 6.3 共享数据分级与策略 5

 6.4 共享伦理要求 5

 6.5 共享安全底线 6

7 共享机制 6

 7.1 共享前数据准备与质量核验 6

 7.2 共享申请 6

 7.3 共享审批 7

 7.4 数据交付 7

8 共享安全与隐私保护 7

 8.1 安全总体框架 7

 8.2 传输与存储安全 8

 8.3 数据脱敏与匿名化 8

 8.4 AI 应用安全与伦理 8

 8.5 隐私保护 8

 8.6 可信数据空间接入 9

9 共享权益与价值分配 9

 9.1 数据权益归属 9

 9.2 价值评估与贡献核算 9

 9.3 收益分配规则 9

附录 A （资料性） 服务水平指标 10

 A.1 服务水平指标清单 10

 A.2 不达标处置 10

附录 B （资料性） 数字合约模板 11

 B.1 数字合约要素 11

 B.2 数字合约记录格式 11

附录 C （资料性） AI 模型卡片模板 12

 C.1 模型卡片要素 12

 C.2 模型卡片记录格式 12

附录 D （资料性） 共享前数据核验清单 13

 D.1 核验项目 13

 D.2 核验记录格式 13

参考文献 14

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。

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专利的责任。

本文件由广东交通职业技术学院提出。

本文件由广东省数字政务协会归口。

本文件起草单位：广东交通职业技术学院、广东省外语艺术职业学院、广东省政务服务数据事务中心、中山大学、广东行政职业学院、广西物流职业技术学院、广州市数字政府运营中心、广州卫生职业技术学院、韩山师范学院、中山市技师学院、揭阳职业技术学院、广东梅州职业技术学院、江门市大数据管理中心、梅州市政务大数据管理中心、潮州市劳动保障信息中心、化州市职业技术学校、公诚管理咨询有限公司、中国联合网络通信有限公司江门市分公司。

本文件主要起草人：

基于人工智能的职业教育教学质量评价数据共享规范

1 范围

本文件规定了基于人工智能技术的职业教育教学质量评价数据共享的总体架构、共享要求与策略、共享机制、共享安全与隐私保护以及共享权益与价值分配。

本文件适用于中等职业学校、高等职业院校及应用型本科院校在教学质量评价过程中，利用人工智能技术对经采集与处理获得的结构化数据开展校内共享、校企共享、跨机构共享，以及向教育主管部门报送、与省市数据平台对接的活动。第三方机构开展职业教育数据共享服务的可参照执行。

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。

GB/T 35273—2020 信息安全技术 个人信息安全规范
GB/T 39786—2021 信息安全技术 信息系统密码应用基本要求
GB/T 41817—2022 信息安全技术 个人信息安全工程指南
GB/T 43697—2024 数据安全技术 数据分类分级规则
GB/T 46347—2025 人工智能 风险管理能力评估
JY/T 0661—2025 教育数据分类分级指南

3 术语和定义

下列术语和定义适用于本文件。

3.1

数据共享 data sharing

将经采集与处理获得的结构化数据，按照约定的权限、格式和接口，提供给经授权的数据使用方使用的过程。

3.2

数据提供方 data provider

在数据共享活动中，拥有数据所有权或合法使用权，并向其他方提供数据的机构或个人。

3.3

数据使用方 data consumer

在数据共享活动中，依法依规申请并使用数据的机构或个人。

3.4

数据沙箱 data sandbox

一种隔离的数据分析环境，数据使用方可在其中对共享数据进行查询和分析，但不直接获取原始数据。

3.5

共享水印 sharing watermark

嵌入共享数据中的不可见标识信息，用于在数据泄露时追溯泄露来源。

3.6

数字合约 digital contract

在共享服务中以机器可执行编码形式存在的协议记录，包含共享数据项、使用目的、使用控制策略、限制条款与违约处置等要素，支持自动解析与执行。

3.7

AI 就绪度 AI readiness

共享数据满足可被人工智能工具直接加载和使用的程度，包括格式标准化程度、元数据完备程度以及质量报告与模型卡片的齐备情况。

3.8

模型卡片 model card

随人工智能模型一同提供的元数据记录，包含模型版本、训练数据范围、评估指标、公平性测试结果、适用场景与禁忌场景、伦理审查结论等信息。

3.9

共享服务平台 sharing service platform

为数据共享提供目录管理、身份认证、数字合约管理、数据交付、使用控制与审计追溯等功能的平台。

3.10

联邦学习 federated learning

一种分布式机器学习范式，多个参与方在不交换原始数据的条件下，通过交换模型梯度或参数协同训练模型的过程。

3.11

可信数据空间 trusted data space

基于共识规则联接多方主体，实现数据资源共享共用的数据流通利用基础设施。

3.12

数据权益 data rights and interests

数据产生方、管理方、质量责任方与使用方在数据共享活动中依法依规享有的权利与相应义务的总称。

3.13

贡献度 contribution degree

数据使用方或参与方对共享数据所支撑的数据产品、服务或研究成果所作贡献的量化度量值。

3.14

数据报送 data reporting

学校按照法律法规、上级教育主管部门要求与上位数据标准，将评价相关数据提交主管部门或上级平台的行为。数据报送是法定义务驱动的数据出口，与共享活动在法律依据与流程上不同。

3.15

公共数据开放 public data opening

公共管理和服务机构按照规定，将公共数据面向社会公开，供符合条件的单位和个人开发利用的活动。公共数据开放适用《广东省公共数据管理办法》的规定，与共享活动在开放条件与范围上不同。

4 缩略语

下列缩略语适用于本文件。

AI：人工智能（Artificial Intelligence）

API：应用程序编程接口（Application Programming Interface）

CSV：逗号分隔值（Comma-Separated Values）

HTTPS：超文本传输安全协议（Hypertext Transfer Protocol Secure）

JSON：JavaScript 对象表示法（JavaScript Object Notation）

NLP：自然语言处理（Natural Language Processing）

SLA：服务水平协议（Service Level Agreement）

SM4：SM4 分组密码算法

5 总体架构

5.1 总体架构

共享体系总体架构自下而上包括共享前准备与核验、共享申请与审批、合约与使用控制、共享交付与审计四层，以及顶部的共享权益与价值分配层，右侧为数据报送与平台对接出口，贯穿全流程的为共享安全、隐私保护与分级管控要求，总体框架见图1。



图1 数据共享总体框架

5.2 共享层级与范式

按数据流动边界，共享活动宜分为三个层级，各层级的适用场景与技术路线见 6.2。

- 校内共享：数据在院校内部流动，不跨越院校边界。
- 校企共享与跨校共享：数据经可信数据空间连接器交换，原始数据留校，对外提供加工产物或服务接口。
- 跨域沙箱共享：数据不离开提供方域，双方在隔离环境中完成联合建模。

6 共享要求与策略

6.1 共享基本原则

数据共享活动遵循下列原则。

- 合法合规：数据共享应符合国家有关数据安全与个人信息保护的规定。涉及个人信息的数据共享，应符合 GB/T 35273—2020 的相关规定。
- 最小必要：共享的数据范围以满足数据使用方的正当使用目的为限，不应超出必要范围。共享前应进行必要性审查，确认每一项共享数据均为达成使用目的所不可或缺。
- 数据不出域、价值可流动：原始数据不离开提供方的可信域。共享的是经处理规范全链路处理后的去标识化画像、聚合视图、技能掌握度矩阵、模型服务接口等加工产物。提供方对原始数据负安全主体责任，对共享出去的加工数据负使用监督责任。
- 使用控制可落地：共享权限的控制应延伸至访问后的使用行为，包括能否导出、能否转发、使用期限长短等。校企共享、跨校共享等场景应通过数字合约实现使用控制的可执行、可审计、可追溯。
- 依法报送与合规开放：向主管部门报送应遵循法定报送义务与上位数据标准的规定，并确定报送数据项、频率与范围；公共数据开放应依法脱敏后开放。二者均不应超出法定目的使用数据。

- f) 全程可追溯：数据共享的全过程——从共享申请、审批、合约签署、数据交付、使用操作到权限回收——应记录完整审计日志，支持从数据使用方反向追溯至数据提供方。

6.2 共享层级与范式

6.2.1 校内共享

院校内部教务部门、教学管理人员、任课教师之间的数据共享，应符合以下要求。

- a) 技术要求：采用基于角色的访问控制与属性控制相结合的权限模型，数据在院校内网流动，不跨越院校边界；审计日志留存于院校本地。
- b) 数据范围：一般数据（L3、L2、L1），包括课堂质量全景视图、学习画像、课程知识图谱等。

6.2.2 校企共享与跨校共享

职业院校与合作企业之间、院校之间的数据共享，应符合以下要求。

- a) 技术要求：宜采用可信数据空间架构，能力要求参考 TDSA/A-001—2025。院校侧部署提供方连接器，原始数据留校，对外提供去标识化画像服务接口；企业侧部署使用方连接器，通过数字合约申请数据访问权限；运营方部署共享服务平台，负责合约协商、使用控制、沙箱管理与审计存证。
- b) 数据范围：一般数据中的较高敏感级别（L3），包括技能掌握度矩阵、实训操作合规率、课堂专注度曲线等。应通过数字合约限定使用目的，不应进行二次传输与反向推断个人身份。
- c) 导出控制：导出任何含学生或班级信息的报表应嵌入共享水印；默认不导出原始查询结果，仅允许导出聚合报表。

6.2.3 跨域沙箱共享

数据使用方需要基于院校数据开展联合建模，且原始数据不离开校园网，应符合以下要求。

- a) 技术要求：采用数据沙箱架构。院校侧沙箱承载学习画像、技能掌握度矩阵等数据，对外提供梯度接口或统计查询接口；使用方侧沙箱承载岗位能力模型、课程标准等数据。
- b) 联合建模：通过联邦学习协议交换模型梯度，不交换原始数据；沙箱内所有操作应记录审计日志。
- c) 数据范围：一般数据与重要数据中的较低敏感部分（L3 至 L4），应经伦理审查后开放。
- d) 模型产出：联合建模产出的模型应附带模型卡片，记录训练数据范围、评估指标、公平性测试结果与伦理审查编号；模型应回传给数据提供方。

6.2.4 向主管部门报送

学校向各级教育行政部门、省级教育科研机构报送教育教学质量评价数据，或接受数据核查、质量评估与抽样复核。

- a) 对外数据交换平台：学校应建立对外数据交换平台，打通学校数据与上级教育主管部门的数据通道。对外数据交换平台的数据抽取、清洗、集成、推送与集成过程监控能力，应符合《职业院校智慧校园规范（试行）》8.3.3.3 的相关规定。
- b) 报送依据：报送的数据项、字段结构与接口规范应符合《职业院校数字基座高职数据标准及接口规范（试行）》与《职业院校数字基座中职数据标准及接口规范（试行）》的规定，本文件不另行规定报送数据项。
- c) 频率与范围：报送频率与报送范围应执行上述上位规范的规定；上位规范未规定的，由学校与主管部门在报送协议中约定并留存记录。报送应支持结构化、半结构化与非结构化数据的定时、定期与实时交换。
- d) 核验与审批：报送前应按附录 D 执行分级核验与去标识化验证。涉及 L4 及以上数据的报送应经双人审批；报送对象限于法定报送渠道，不应通过互联网直接报送。
- e) 数据范围：按上位规范确定的数据子集与数据子类范围。评价数据的扩展数据子类应按 6.3 的要求另行编码，并向数据主管部门申请。
- f) 责任分工：学校对报送数据的真实性、准确性、完整性负责；接收方对数据的使用范围与用途负责，不应超出评价与监管目的使用。

6.2.5 与省市数据平台对接

学校将教育教学质量评价数据接入省级数据资源“一网共享”平台，或经授权向公共数据开放平台推送。

- a) 接入通道：与省市数据平台的对接应通过政务外网或经批准的安全通道完成，不应通过互联网直接交换。
- b) 数据资源登记：接入前应完成公共数据资源登记。登记的申请、受理、审核、公示与异议处理，按属地公共数据资源登记管理实施政策要求的流程执行。
- c) 目录编制：数据资源目录应按“一数一源”原则编制，明确数据责任部门、来源、更新机制、质量标准与使用方法。同一数据项不应在多个部门重复采集。
- d) 开放与脱敏：推送至公共数据开放平台的数据应按属地公共数据管理政策的要求依法脱敏后开放；涉及个人隐私、商业秘密或者依据法律、法规、规章及国家规定不得开放的，不予开放。
- e) 服务与记录：依托平台的可用性与接口稳定性应满足附录 A 的服务水平要求。接入记录、登记结果与推送记录应留存，保存期限与数字合约期限一致。
- f) 数据范围：经登记与批准的数据资源。涉及个人信息的数据应完成去标识化或匿名化；L4 及以上数据不应接入省市数据平台。

6.3 共享数据分级与策略

共享数据分级应符合 JY/T 0661—2025 与 GB/T 43697—2024 的相关规定。当同一数据项同时属于多个类别时，按就高从严原则确定共享策略。各级数据按共享场景确定的策略见表3。

表 3 共享数据分级与策略

级别	名称	覆盖数据示例	校内与机构共享	向主管部门报送	接入省市数据平台
L5	核心数据	全国与全省职教基础库、统一身份认证根数据	不共享	按国家平台统一要求报送，报送范围与方式由国家平台规定	不接入
L4	重要数据	全校学生实名成绩单、含人脸或声纹的原始采集数据	仅校内共享，需双人审批	按 6.2.4 执行，双人审批，出校范围限于报送对象	不接入
L3	一般数据（较高）	去标识化学习画像、技能掌握度矩阵、课堂聚合视图	机构共享主力：去标识化、数字合约、使用控制	可按 6.2.4 报送	经登记与批准后可接入
L2	一般数据（中等）	课程画像、专业画像、教师教学画像、增值评价结果	可共享，聚合口径下提供	可按 6.2.4 报送	经登记与脱敏后可接入
L1	一般数据（较低）	课程公开信息、公开统计指标	可共享，常规审批	可按 6.2.4 报送	经登记后可接入

注1：各级别的具体数据项应依据 JY/T 0661—2025 的分类分级规则确定，并列入共享数据目录。

注2：向主管部门报送的，同时执行 6.2.4 的规定；接入省市数据平台的，同时执行 6.2.5 的规定。数据进入任何一种数据出口前，均应按附录 D 执行分级核验，核验未通过的数据不应进入相应流程。

注3：向主管部门报送是履行法定报送义务，不因本表策略而免除；但报送范围应与上位规范确定的数据子集一致，不应超出报送对象法定职责所需。

6.4 共享伦理要求

共享活动受人工智能伦理要求的约束，并且符合下列要求。

- a) 中性标签：共享数据中不应包含“差生”“问题学生”“落后生”等负面定性标签，仅输出中性、建设性的描述。
- b) 非歧视性内容：共享数据不应包含任何形式的性别、地域、民族、家庭背景等方面的歧视性表述。
- c) 可解释性随行：共享的人工智能评价结果应附带可解释性说明，说明其数据依据与判断逻辑。

- d) 人工监督保留：共享的人工智能评价结果仅供数据使用方参考，涉及学生重大利益的决策应经过人工审核确认。
- e) 伦理审查前置：涉及 L3 级以上数据或可能影响学生重大利益的共享活动，应在共享申请审批前通过伦理审查。

6.5 共享安全底线

数据共享活动满足下列安全底线要求。

- a) 传输加密：所有共享接口应采用 HTTPS 或 SM4 加密协议，不应使用明文传输或已废弃的协议版本。
- b) 密钥管理：接口鉴权密钥与加密密钥应由密钥管理系统统一管理，不应硬编码。
- c) 共享水印：所有导出的共享数据与报表应嵌入不可见水印，水印内容至少包含数据使用方标识、共享时间与数字合约编号。
- d) 最小权限：每个数据使用方仅应获得完成其使用目的所需的最小数据集与最小操作权限。
- e) 数据出境管控：涉及跨境数据共享的，应按照国家关于数据出境安全管理的相关规定开展安全评估。职业教育教学质量评价数据原则上不出境。

7 共享机制

7.1 共享前数据准备与质量核验

共享前应完成数据准备与质量核验，并符合下列要求。

- a) 来源确认：共享的数据项应具备可追溯的处理链路记录。不应共享来源不明或未经处理的数据。
- b) 质量核验：共享前应确认质量等级达到所在共享层级的要求。核验未通过的数据不应进入共享流程。
- c) 脱敏与封装：应按 8.3 完成去标识化与泛化处理，并按约定的数据格式与接口规范完成封装。封装记录应包含数据项、字段结构、更新频率与责任方。
- d) AI 就绪度评估：拟供人工智能工具使用的共享数据，应评估其 AI 就绪度，评估项目见附录 D。评估结果应随数据一并提供。
- e) 分级核验：应按 6.3 确认数据分级与共享策略匹配，并列入共享数据目录。核验项目见附录 D。
- f) 核验记录：核验过程应形成记录，记录核验时间、核验人、核验项目与核验结论，并归档留存。核验记录格式见附录 D。

7.2 共享申请

数据使用方发起共享申请前完成下列准备。

- a) 身份注册：应在共享服务平台或可信数据空间运营方完成机构身份注册，获取唯一机构标识与接入凭证。注册信息应包含机构名称、统一社会信用代码、数据安全负责人信息与网络安全等级保护定级信息。
- b) 能力声明：应声明本机构的数据安全保护能力，包括数据管理制度建设、数据安全技术防护与事件应急响应能力。
- c) 人工智能能力声明：拟使用人工智能模型处理共享数据的，应声明模型类型、用途、训练框架与安全防护措施。
- d) 目的声明：应明确声明所需共享数据的使用目的，使用目的应具体、合法并与职业教育教学质量评价直接相关。
- e) 共享申请内容应包括下列信息：
 - 1) 需求方信息：机构名称、机构标识、联系人与联系方式。
 - 2) 所需数据项：从共享数据目录中选择，可指定数据切片范围。切片应支持按时间维度（学年、学期、月份）、组织维度（专业、年级、班级）、课程维度（课程标识、知识点标识、技能点标识）与数据维度（数据项名称、AI 就绪度）组合筛选。

- 3) 使用目的与方式：具体描述业务场景与预期产出；明确使用方式，包括查询、统计分析、模型训练、导出报表等。
- 4) 使用期限：申请数据使用的起止时间。期限上限由数字合约约定。
- f) 安全与人工智能使用承诺：承诺遵守数据安全保护要求、不超范围使用、不进行二次传输、不反向推断个人身份；涉及人工智能模型训练的，应承诺不训练可能对学生产生歧视性影响的模型，并将模型回传给数据提供方。
- g) 申请提交方式与受理应符合下列要求：
 - 1) 提交方式：可通过共享服务平台门户提交，或通过标准化接口自动提交。
 - 2) 受理确认：系统应生成唯一申请编号并返回受理确认。

7.3 共享审批

共享审批包括人工智能辅助预审与人工审批两个环节。

- a) 人工智能辅助预审：共享服务平台可利用人工智能技术对申请进行自动预审，预审内容包括：身份核验（利用知识图谱与实体链接技术核验机构标识真实性）；数据项可用性（确认所申请数据项处于可用状态且属于该使用方的允许范围）；使用目的合规性（利用自然语言处理技术对使用目的作语义分析，判断是否属于允许范围）；使用期限合理性；风险评分（基于历史共享记录、安全能力等级与数据敏感度计算风险分值）。
- b) 预审处置：预审不通过的应自动退回申请并说明原因；风险评分高于阈值的应标记为需重点审查；预审通过的进入人工审批环节。
- c) 人工审批：数据提供方的数据管理员或授权审批人对申请进行审批。审批期限应在平台服务说明中公布。
- d) 伦理审查：涉及 L3 级以上数据或可能影响学生重大利益的共享活动，应在审批前通过伦理审查。
- e) 审批结论：审批结论分为批准、有条件批准与拒绝三种。有条件批准的，应说明需补充的材料或需调整的使用范围；拒绝的，应说明拒绝理由。

注：审批过程与结论应记录并留存，审计日志应支持从审批结论追溯至预审过程与申请材料。

7.4 数据交付

数据交付符合下列要求。

- a) 交付方式：可采用接口交付、文件交付或沙箱内查询方式。交付方式应在数字合约中约定。
- b) 智能缓存：对高频查询的共享数据项，可采用缓存策略降低查询延迟。缓存位置与缓存期限应符合 8.2 的规定。
- c) 格式适配：可根据数据使用方的工具类型适配数据格式，可交付 JSON、CSV 或约定格式，减少使用方的格式转换成本。格式适配不应改变数据项的业务含义。
- d) 切片推荐：数据量较大时，可根据使用方的历史使用模式推荐数据切片策略，提高查询效率。
- e) 交付日志：每次交付应记录合约编号、交付时间、交付方式、交付数据项、数据量、AI 就绪度等级、交付状态与失败原因。
- f) 服务水平：交付服务水平应满足附录 A 的要求。附录 A 中的指标值为参考值，具体目标值由运营方与数据使用方在数字合约中约定。

8 共享安全与隐私保护

8.1 安全总体框架

安全责任边界应符合下列规定。

- a) 数据提供方：负责原始数据的安全存储与合规处理，负责共享前数据的脱敏与质量验证。
- b) 共享服务平台运营方：负责平台自身的安全防护、访问控制、传输加密、审计日志、安全监控与应急响应。
- c) 数据使用方：负责遵守数字合约约定的使用控制策略，负责其内部系统对共享数据的安全存储与合规使用，负责人工智能模型训练过程中的安全防护。

- d) 审计方：负责对共享活动进行独立审计，发现安全问题及时通报。

8.2 传输与存储安全

传输与存储安全满足以下要求。

- a) 加密协议：共享接口应采用 HTTPS（TLS 1.3）或 SM4 加密协议，不应使用明文传输或已废弃的协议版本。
- b) 双向认证：共享服务平台与数据提供方、数据使用方之间应进行双向证书认证，防止中间人攻击。
- c) 密钥管理：鉴权密钥与传输加密密钥应由密钥管理系统统一管理，更换周期由运营方依据 GB/T 39786—2021 的相关规定确定，不应硬编码在配置文件中。
- d) 加密存储：共享服务平台存储的共享数据与缓存数据应采用合规的加密算法加密存储。
- e) 密钥隔离：加密密钥与加密数据应隔离存储，由专用密钥管理系统管理，访问密钥需双重授权。
- f) 临时数据清理：查询结果、沙箱中间结果等临时数据在使用完毕后应立即清除，最长保留期限由数字合约约定。
- g) 冷数据加密：归档的审计日志、历史合约等冷数据同样应加密存储，解密密钥需单独保管。

8.3 数据脱敏与匿名化

共享数据在纳入目录前应完成脱敏处理，去标识化处理应符合 GB/T 35273—2020 的相关规定，并符合下列要求。

- a) 直接标识符替换：学生姓名、学号、身份证号等直接标识符应替换为匿名标识。
- b) 准标识符泛化：对出生年月、籍贯、家庭住址等准标识符应进行泛化处理。
- c) 生物特征处理：人脸图像、声纹特征等生物特征不应纳入共享数据，仅共享基于这些特征分析得出的聚合统计指标。
- d) 脱敏验证：共享数据输出前应经过自动脱敏验证工具检查，确保不包含可反溯至个人的信息。验证项目见附录 D。

8.4 AI 应用安全与伦理

共享数据涉及人工智能模型时，应符合 GB/T 46347—2025 的相关规定，并符合下列要求。

- a) 模型加密：共享的模型文件应采用合规的加密算法加密传输和存储。
- b) 模型签名：模型文件应附带数字签名，数据使用方加载前验证签名，防止模型被篡改或植入后门。
- c) 模型卡片随行：模型共享应附带模型卡片，卡片要素见附录 C 表 C.1。
- d) 模型溯源：模型的血缘信息应可追溯至其训练所使用的数字合约，确保模型使用不超出原始数据的使用目的。
- e) 模型回传：模型使用方应在约定期限内向数据提供方回传模型或其评估结果。
- f) 验收测试：共享或联合建模产出的模型，数据提供方应开展验收测试，并按本文件附录 B 表 B.5 的格式记录测试结果。未经验收测试的模型不应进入生产环境。

8.5 隐私保护

共享活动中的个人信息保护应符合 GB/T 35273—2020、GB/T 41817—2022 的相关规定，并符合下列要求。

- a) 最小化共享：共享的数据应以聚合统计与去标识化画像为主，不应共享包含个人身份信息的原始数据。
- b) 知情同意验证：共享申请审批时，应验证数据提供方已取得学生或监护人的知情同意。
- c) 用途限定：共享数据仅限用于数字合约约定的使用目的，不应用于商业营销、学生排名公开、招聘筛选等用途。
- d) 数据主体权利：学生有权查询其被共享了哪些数据、共享给了谁、用于什么目的。共享服务平台应提供相应的查询接口。
- e) 影响评估：共享范围、数据类型或使用目的发生重大变更时，应重新开展个人信息安全影响评估。

8.6 可信数据空间接入

采用可信数据空间架构的共享场景，应满足下列要求。

- a) 框架符合：可信数据空间的能力构成应符合 TDSA/A-001—2025 确定的框架。
- b) 可信管控：应提供接入认证与身份核验能力，并对接入的数据资源、数据产品与数据服务开展规范性审查。
- c) 日志存证：应具备追踪与记录数据来源、流动路径与使用状态的能力，保证操作行为记录真实且不可篡改。
- d) 互联互通：宜按照目录标识、身份认证与接口要求实现各类数据空间互联互通。

9 共享权益与价值分配

9.1 数据权益归属

数据权益归属应按管用分离原则明确各方责任与权利，并符合下列要求。

- a) 生产单位：对数据采集的完整性与真实性负责。
- b) 管理单位：对数据标准执行、目录管理与共享策略制定负责。
- c) 质量责任单位：对数据质量符合规范要求负责，并负责组织质量问题处置。
- d) 使用单位：按数字合同约定的使用目的与使用控制策略使用数据，不超范围使用、不进行二次传输、不反向推断个人身份。
- e) 权益登记：各方权益与责任应登记于共享数据目录，并在数据提供方与使用方之间留存记录。权益登记项见附录 D。

9.2 价值评估与贡献核算

价值评估与贡献核算符合下列要求。

- a) 贡献度记录：数据使用方基于共享数据形成数据产品、服务或研究成果的，可记录其贡献度。贡献度应记录贡献类型、数据使用量、产出形式与产出规模，记录格式见附录 B 表 B.2。
- b) 数据依据：贡献度的计算应以下列数据为依据：合约编号、数据使用记录、交付日志、使用行为审计记录与产出登记信息。上述数据均应可追溯至数字合约。
- c) 计算方法：贡献度的计算方法与权重由参与方在数字合约中约定，并记录约定内容。计算方法变更的，应保留变更记录。
- d) 价值评估：可将贡献度作为价值评估的参考依据之一。价值评估还应考虑数据稀缺性、时效性与替代成本等因素，评估因素由参与方约定。
- e) 留痕要求：贡献度记录应与合约记录关联留存，保存期限与合约期限一致。合约终止后，记录的保存期限由参与方约定。

9.3 收益分配规则

收益分配规则符合下列要求。

- a) 规则约定：分配规则应由参与方在数字合约中约定，约定内容应包含分配依据、计算方式、支付安排与争议处理。本文件不规定具体分配比例。
- b) 规则记录：分配规则应作为数字合约的组成部分留存，并记录规则的版本与生效时间。规则变更的，应重新签署或补充约定。
- c) 规则公示：分配规则及其执行情况应在参与方范围内公示。涉及多方的，由共享服务平台提供公示支持。公示应采用可访问的方式，并留存公示记录。
- d) 执行记录：分配执行情况应形成记录，记录时间、金额或资源、依据的规则版本与经办人，并与贡献度记录关联。
- e) 争议处理：对分配结果存在争议的，应按合同约定的争议处理机制解决；协商不成的，可按合同约定申请第三方评估或仲裁。

附 录 A
(资料性)
服务水平指标

A.1 服务水平指标清单

共享服务的服务水平指标见表A.1。

表 A.1 服务水平指标清单

指标名称	定义	参考值	测量方式	统计周期
接口可用率	接口正常响应次数占总请求次数的比例	99.5%	平台监控系统统计	月度
接口响应时间(P50)	所有请求响应时间的第 50 百分位值	500 ms	平台监控系统统计	月度
接口响应时间(P99)	所有请求响应时间的第 99 百分位值	2000 ms	平台监控系统统计	月度
审批完成率	在承诺时限内完成审批的申请数占申请总数的比例	95%	审批记录统计	月度
合约履约率	按约完成交付与响应请求的次数占应履约次数的比例	99%	合约记录统计	季度
并发支持能力	同时处理的有效请求数	由运营方公布	平台监控系统统计	—
调用频次上限	单位时间内允许的接口调用次数	由合同约定	平台监控系统统计	—

注1：表中数值为参考值。接口可用率、响应时间、审批完成率与合约履约率的具体目标值，由运营方与数据使用方在数字合约中约定。

注2：并发支持能力与调用频次上限属平台能力指标，由运营方公布或由数字合约约定，本文件不作统一规定。

注3：本附录对应本文件 7.4，服务水平不达标的处置见 A.2。

A.2 不达标处置

服务水平指标未达到合约约定目标值时，应采取下列处置方式。

- a) 轻微偏离：接口可用率或合约履约率低于约定值但高于兜底值的，运营方应向受影响的数据使用方说明情况，并在下一统计周期内完成问题修复。
- b) 显著偏离：低于兜底值的，运营方应提交事故分析报告，并按数字合约约定承担相应责任。
- c) 审批时限偏离：审批完成率低于约定值的，应优化审批流程或调整审批人员配置。
- d) 申诉处理：数据使用方对处置结果有异议的，可按数字合约约定的争议处理机制提出申诉。

注：具体兜底值与责任承担方式由数字合约约定，本文件不作统一规定。

附 录 B
(资料性)
数字合约模板

B.1 数字合约要素

数字合约要素见表B.1。

表 B.1 数字合约要素

要素	说明
contractId	合约唯一标识
providerOrgId	数据提供方机构标识
consumerOrgId	数据使用方机构标识
dataItems	共享数据项清单，含数据项编码、分级级别与 AI 就绪度
purpose	使用目的，应具体、合法且与教育教学质量评价直接相关
usageMode	使用方式：查询、统计分析、模型训练、导出报表等
usagePeriod	使用期限，含起始时间与截止时间
controlPolicy	使用控制策略：导出限制、转发限制、截屏限制等
prohibitedTerms	限制条款：不含二次传输、不含反向推断个人身份等
algorithmPolicy	人工智能使用约定：模型类型、训练范围、回传要求
contributionRule	贡献度计算方法与权重约定
allocationRule	收益分配规则的引用与版本号
slaTargets	服务水平目标值，引用附录 A 表 A.1 的指标项
breachHandling	违约处置与责任承担方式
signature	签署记录，含签署时间与签署方标识

注：合约要素的必填项由运营方根据共享层级确定，涉及 L3 级以上数据的合约应完整包含全部要素。

B.2 数字合约记录格式

贡献度记录作为数字合约的关联记录，其字段见表B.2。

表 B.2 贡献度记录字段

字段	类型	说明
contractId	string	关联的合约唯一标识
contributionType	string	贡献类型：数据产品、服务、研究成果等
dataUsageVolume	number	数据使用量，记录数或调用次数
outputForm	string	产出形式：数据集、接口服务、报告、模型等
outputScale	string	产出规模描述
methodVersion	string	贡献度计算方法版本，与合约约定一致
contributionValue	number	按约定方法计算的贡献度值
evidenceRef	string	依据的交付日志与审计记录引用
recordedAt	string	记录时间，UTC ISO 8601 格式
recordedBy	string	记录人标识

附录 C
(资料性)
AI 模型卡片模板

C.1 模型卡片要素

模型卡片要素见表C.1。

表 C.1 模型卡片要素

要素	必填	说明
modelName	是	模型名称
modelVersion	是	模型版本号，格式为 v{主版本}.{次版本}
modelType	是	模型类型：分类、回归、聚类、序列标注、生成等
trainingDataScope	是	训练数据范围，含使用的合约编号
evaluationMetrics	是	评估指标与结果
fairnessTestResult	是	公平性测试结果，含群体间性能差异
applicableScenarios	是	适用场景
prohibitedScenarios	是	禁忌场景
ethicsReviewId	是	伦理审查编号
providerOrgId	是	模型提供方机构标识
providerContact	否	模型提供方联系人与联系方式

注：本附录对应本文件 8.4。模型卡片随模型一并提供，卡片内容应可核验，其血缘信息应可追溯至训练所使用的合约。

C.2 模型卡片记录格式

模型卡片记录格式见表C.2。

表 C.2 模型卡片记录字段

字段	类型	说明
modelName	string	模型名称
modelVersion	string	模型版本号
modelType	string	模型类型
trainingDataScope	array	训练数据范围，含合约编号列表
evaluationMetrics	object	评估指标与结果
fairnessTest	object	公平性测试结果，含测试维度、对比组与指标
applicableScenarios	array	适用场景列表
prohibitedScenarios	array	禁忌场景列表
ethicsReviewId	string	伦理审查编号
providerOrgId	string	模型提供方机构标识
issuedAt	string	签发时间，UTC ISO 8601 格式
signature	string	模型文件数字签名值

附录 D
(资料性)
共享前数据核验清单

D.1 核验项目

共享前核验项目见表D.1。核验未通过的数据不应进入共享流程。

表 D.1 共享前核验项目

类别	核验项目	判定要求
来源	来源可追溯性	数据项来自处理规范输出目录，具备处理链路记录
来源	处理链路完整性	清洗、标注、融合各环节记录完整、可追溯
质量	质量等级	达到所在共享层级要求的等级
质量	格式规范性	字段结构符合约定格式，元数据齐备
分级	分级准确性	分级级别与 JY/T 0661—2025 判定一致，就高从严执行
分级	策略匹配性	共享策略与分级级别对应，未超范围共享
脱敏	直接标识符替换	姓名、学号等已替换为匿名标识
脱敏	准标识符泛化	出生年月、籍贯等已泛化处理
脱敏	生物特征排除	不含人脸图像、声纹特征等生物特征原始数据
脱敏	反推风险验证	通过自动验证工具检查，不可反溯至个人
AI 就绪度	格式标准化	格式可被人工智能工具直接加载
AI 就绪度	元数据完备性	含字段含义、单位、更新频率与责任方
AI 就绪度	质量报告齐备	附质量报告，标明指标项与实测值
AI 就绪度	模型卡片齐备	涉及训练的，附模型卡片
权益	权益登记	各方责任与权益已登记于共享数据目录
权益	责任约定	数据提供方与使用方责任边界已明确

D.2 核验记录格式

核验记录格式见表D.2。

表 D.2 共享前核验记录字段

字段	类型	说明
dataItemCode	string	数据项编码
gradeLevel	string	分级级别：L1~L5
checkTime	string	核验时间，UTC ISO 8601 格式
checker	string	核验人标识
checkItems	array	核验项目及其判定结果
aiReadiness	string	AI 就绪度：满足、基本满足、不满足
conclusion	string	核验结论：通过、不通过、限期整改
rectifyDeadline	string	整改期限，未通过且可整改时填写
attachedFiles	array	质量报告、模型卡片等附件引用

注：核验记录应归档留存，保存期限与对应数字合约期限一致。

参 考 文 献

- [1] GB/T 35273—2020 信息安全技术 个人信息安全规范
 - [2] GB/T 43697—2024 数据安全技术 数据分类分级规则
 - [3] JY/T 0661—2025 教育数据分类分级指南
 - [4] TDSA/A-001—2025 可信数据空间能力要求
 - [5] ISO/IEC DIS 26450 Information technology — Eclipse dataspace protocol (DSP)
 - [6] ISO/IEC DIS 26451 Information technology — Eclipse decentralized claims protocol (DCP)
 - [7] 国家数据局. 可信数据空间发展行动计划（2024—2028年）
 - [8] 《职业院校数字基座高职数据标准及接口规范（试行）》
 - [9] 《职业院校数字基座中职数据标准及接口规范（试行）》
 - [10] 《职业院校智慧校园规范（试行）》
 - [11] 《职业院校人工智能应用指引（2026版）》
 - [12] 《广东省公共数据管理办法》
 - [13] 中共广东省委办公厅、广东省人民政府办公厅. 关于加快公共数据资源开发利用的实施意见
 - [14] 广东省政务服务和数据管理局. 广东省公共数据资源登记管理实施细则（试行）
-