

团体标准《标准开源格式规范》（征求意见稿）编制说明

广东财经大学

广东省节能减排标准化促进会

二〇二六年九月

一、工作简况包括任务来源、主要工作过程、主要参加单位所做工作

1. 任务来源

2026 年，大语言模型（Large Language Model，LLM）基于 Transformer 架构的自注意力机制，聚焦于整合文本与视觉的端到端架构，并融入强化学习，赋予类人的慢思考与推理能力，将百万级超长上下文能力进行开源普惠化。Agent 能够连续自主编程 60 分钟以上，无需人工干预即可完成数据库设计、安卓模拟器开发等复杂工程任务，具备出色的长程规划与自我迭代能力。特别是 OpenClaw 从“传统聊天机器人”动口角色到“数字员工”动手角色的突破，以及各地扶持政策的落地，Agent 已到了迈向“规模化落地”的关键时期。基于大语言模型，沿着“LLM → AI ChatBot → Prompt → Context → Harness → Loop”技术发展跃进路线，以及“Token”、“API”、“MCP”、“CLI”、“Skill”等词涌现，代表着 AI 工程化从“调教模型”到“设计系统”的演进路径。

在搭建的标准开源社区（http://www.perfectquality.cn/）的基础上，围绕核心标准，定位于标准实际应用，基于现有大语言模型应用的技术环境，需要建设标准智能体（Standard Agent）。标准规范的内容，需要基于现有大语言模型应用的技术环境，采用 Markdown 格式来标识标准内容，实现机器可显示标准文档、机器可读取标准文档、机器可读取标准内容、机器可执行标准内容、机器可解释标准内容、机器可更新标准内容。标准应用的资源，主要包括知识库、开源代码和元数据格式发布，统一规范标准应用资源。标准参考的案例，

主要包括数据集及数据库和典型案例，示范推广标准实际应用。标准更新的机制，依据大语言模型 Markdown 格式的更新机制，借助 Agent 现有技术底座实现标准自主更新。

目前，标准文档主要以 PDF、Word 等非结构化格式发布，机器可读性差，AI 智能体难以直接解析、检索和执行标准内容。虽然 Google Cloud 于 2026 年 6 月 12 日发布了开放知识格式 OKF(Open Knowledge Format)，提供了一套基于“文件夹+Markdown”的文件规范，但该规范面向通用知识库，尚未针对标准文档的元数据、结构层次和交换需求进行专门化设计。为此，广东财经大学提出制定《标准开源格式规范》团体标准，以建立标准开源格式 SOSF（Standard Open Source Format），实现人类可读、智能体可解析、跨平台可交换的标准知识库规范。经广东省节能减排标准化促进会评审，标准符合立项条件，于 2026 年批准立项。

2.主要工作过程

在本标准制定计划下达后，广东财经大学、广东省节能减排标准化促进会、广东质高标准化服务有限公司等单位共同成立了标准编制小组，开展了具体工作。

1）立项申请与调研阶段

2026 年 9 月，召开了项目启动会，并在前期标准开源社区建设的基础上，广东财经大学、广东省节能减排标准化促进会、广东质高标准化服务有限公司达成共识，正式成立标准编制工作小组，开始收集、分析、总结国内外有关标准资料和文献，提出相应的编写要求。同时召开工作组第一次工作会议。

2）标准起草阶段

标准立项后，由上述单位组建标准起草工作组，启动标准草案的编制工作。起草工作组在系统梳理标准文献元数据、知识库格式、大语言模型应用等技术资料的基础上，参考 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》、GB/T 22373—2021《标准文献元数据》等现行标准，结合标准开源社区实际应用经验，以及 Google Cloud 发布的开放知识格式 OKF，按照 GB/T 1.1—2020 的要求和规定编写内容，确定了标准开源格式的标准包结构、标准概念文件、头部元数据、正文等技术内容，形成了标准草案。

3）征求意见阶段

本标准现处于征求意见阶段。项目组已将标准征求意见稿面向社会公开征求意见，广泛征集行业专家、相关企业及科研机构的反馈建议，以进一步完善标准技术内容。

3. 主要参加单位和工作组成员及其所做工作

1）主要参加单位

本标准由广东财经大学提出。归口单位为广东省节能减排标准化促进会。

广东财经大学以经济学、管理学、法学为主体，多学科协同发展，现有应用经济学、国家安全学 2 个博士学位授权一级学科点，11 个硕士学位授权一级学科点、23 个硕士专业学位授权点，18 个国家级一流本科专业建设点。拥有全国首批经管类国家级实验教学示范中心、国家级虚拟仿真实验教学中心等平台，社会科学总论、工程学进入 ESI 全球前 1%。学校立足粤港澳大湾区，坚持“求真、致用”，正加快建设大湾区一流财经大学。

2）工作组成员及分工

标准主要起草人团队涵盖标准化、人工智能、知识管理、软件开发等领域的专业人员。主要工作分工如下：

工作内容	责任单位/人员
标准文本编写	广东财经大学、 广东省节能减排标准化促进会
标准开源格式技术方案确认	广东财经大学、 广东质高标准化服务有限公司
标准智能体应用验证	广东财经大学、 广东质高标准化服务有限公司
标准归口管理与组织评审	广东省节能减排标准化促进会

二、团体标准编制原则和确定团体标准主要内容的论据

1. 编制原则

1）规范性原则。本标准的制定严格按照 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》的要求和规定，编写本标准内容。标准文本简明扼要，能直接引用的标准尽量引用。标准的文字表述准确、简明、易懂、结构合理、层次分明、逻辑严谨。

2）协调性原则。本标准在编写过程中遵循国家标准的方针政策和法规制度，与其他国家标准、行业标准相协调一致。本标准在技术内容上与 GB/T 22373—2021《标准文献元数据》保持协调，标准概念文件的头部元数据字段选取参考 GB/T 22373—2021 中的 Number、

Name、Announcing_body、publishing_date、effective_date、ICS_number、CCS_number、industry、drafting_body、drafting_person、scope、patent 等字段。

3）开放性原则。SOSF 是一种开放的、对人类和智能体都友好的格式，用于表示标准。该格式刻意保持极简：一个包含带 YAML 头部元数据的 Markdown 文件的目录。没有模式注册表，没有中央权威机构，也不需要专用工具。

4）可读性与可解析性原则。标准开源格式应实现人类可读，无需工具即可阅读；智能体可解析，无需定制 SDK；可在版本控制中对比差异；可跨工具、跨组织、跨时间移植。

2.编制依据

本文件在技术指标和试验方法方面参考以下国家标准：

GB/T 22373—2021《标准文献元数据》

Google Cloud 开放知识格式 OKF（2026 年 6 月 12 日发布）

标准开源社区（http://www.perfectquality.cn/）建设实践

三、主要内容

标准共包括 5 个章节内容：

1) 第一章节为范围：本文件给出了标准开源格式的标准包结构和标准概念文件等内容。本文件适用于基于大语言模型的AI智能体标准知识库建设，其他文档类的AI智能体知识库建设也可参照执行。

2) 第二章节为规范性引用文件。

3）第三章节为术语和定义。

4）第四章节为标准包结构： 给出了包结构和标准包结构。包是 Markdown 文件的目录树，目录结构与领域无关，生成者可以按照对所捕获知识有意义的任何方式组织概念。包可以以下列方式分发：Git 仓库（推荐，因为它提供历史记录、归属和差异对比）；目录的 tarball 或 zip 归档；更大仓库中的子目录。保留文件名：index.md 目录列表文件和 log.md 更新历史文件不得用于概念文档。

5）第五章节为标准概念文件：给出了标准概念文件的头部元数据和正文要求。头部元数据采用 YAML frontmatter 格式，必填字段为 type，可选字段包括 title、description、resource、tags、generated、verified、status、computation 等。标准概念文件的头部元数据 type 取值为“Document”，title 取值为“standard”，tags 标签参见 GB/T 22373—2021 选取 Number、Name、Announcing_body、publishing_date、effective_date、ICS_number 、 CCS_number 、 industry 、 drafting_body 、drafting_person、scope、patent 等字段。正文在概念文件正文的基础上，规定标准的封面、目次、前言、引言、范围、规范性引用文件、术语和定义、其他章、参考文献、索引、附录等部分为一级标题，相应其下的子标题为二级标题，最多设置三级标题，其余设置为正文。

四、试验验证的分析

本标准为标准文档的数字化、机器可读化提供科学、规范、可操作的格式规范，明确标准包结构、标准概念文件、头部元数据字段及正文结构要求。本标准以 Markdown 和 YAML 为技术底座，实现标准文档的人类可读、智能体可解析、跨平台可交换。

本标准在标准开源社区（http://www.perfectquality.cn/）进行了初步应用验证。社区依据标准开源格式 SOSF 制作和发布 Markdown 格式标准数据，实现标准数据开源。然后依据大语言模型，借助 Agent 现有技术底座，开发建设标准应用的 AI 智能体。标准 AI 智能体的开发建设计划包括:零碳工厂智能体、零碳园区智能体、VCER 碳信用资产智能体等。SOSF 兼顾人类可读与 AI 可自动解析，轻量化，适合 Agent 长期知识库落地。初步验证结果表明，采用 SOSF 格式的标准文档能够被大语言模型直接读取、解析和引用，显著降低了标准智能体开发中的文档预处理成本。

本标准可作为标准知识库建设、标准智能体开发、标准数据交换等技术依据，支撑标准从“纸质文档”走向“机器可执行知识”。预期的经济效果包括：降低标准数字化和智能体开发成本，提升标准内容的可检索性和可复用性；促进标准在 AI 智能体中的规模化应用，推动零碳工厂、零碳园区、VCER 碳信用资产等领域的标准智能化落地；为跨组织、跨平台的标准知识交换提供统一格式，支撑标准开源生态建设。

五、采用国际标准和国外先进标准的程度，以及与国际、国外同类标准水平的对比情况，或与测试的国外样品、样机的有关数据对比情况。

本标准的部分技术内容参考了 Google Cloud 于 2026 年 6 月 12 日发布的开放知识格式 OKF，但没有对国际国外标准进行直接采用。

OKF 是一套基于“文件夹+Markdown”的文件规范，通过定义一种通用标准的文档类格式，生成者（人类、智能体、导出管道）可以写入，指导使用者（智能体、用户界面、搜索索引、确定性代码）应如何阅读和遍历它，促进跨系统和组织的知识交换。本标准在 OKF 的基础上，针对标准文档的特点，增加了标准包结构、标准概念文件头部元数据字段、标准正文层次结构等专门化要求，形成了 SOSF 标准开源格式。

六、产业化情况、推广应用论证和预期达到的经济效果、社会效益和生态效益

1、产业化情况

1）技术成熟度与产业基础

SOSF 参考 Google Cloud 于 2026 年发布的开放知识格式 OKF，结合标准文献元数据要求（GB/T 22373—2021）进行专门化设计。其核心采用 Markdown、YAML、Git 等成熟技术，无需专用工具和模式注册表，具备“cat 可读、git clone 可分发的”极简特性。标准开源社区（http://www.perfectquality.cn/）自搭建以来，已围绕核心标准开展 Markdown 格式标准数据发布和标准智能体建设实践，为本标准的产业化应用奠定了基础。

2）市场空间

在双碳目标、绿色低碳转型和 AI Agent 规模化落地的背景下，零碳工厂、零碳园区、VCER 碳信用资产、产品碳足迹、水足迹、碳减排量核算等领域对标准智能体的需求持续增长。SOSF 作为标准知识库的底层格式规范，可广泛应用于标准智能体开发、标准知识库建设、标准数据交换、标准合规审查、标准教育培训等场景，市场空间广阔。

2、推广应用论证

1）必要性

现有标准文档格式不利于 AI 智能体直接解析和执行，标准知识难以在跨组织、跨平台、跨系统之间高效交换。制定 SOSF 标准，有助于解决标准数字化过程中的格式不统一、元数据不一致、结构不规范等问题，推动标准从“人读”走向“人机共读”，支撑标准智能体规模化落地。

2）技术可行性

SOSF 采用 Markdown 和 YAML，技术成熟、工具链丰富、大语言模型原生支持。标准包以目录树和 Markdown 文件为基本单位，支持 Git 版本管理、差异对比和协作开发。标准概念文件通过 YAML 头部元数据标识标准号、名称、发布机构、发布日期、实施日期、ICS、CCS、起草单位、起草人、范围、专利等信息，结构清晰、易于扩展。标准开源社区的初步应用表明，SOSF 格式可被 AI 智能体直接读取和引用，显著降低知识库构建成本。

3）经济可行性

本标准为推荐性团体标准，采用开源格式，不增加额外授权成本。标准包可通过 Git 仓库、zip 压缩包或子目录方式分发，边际传播成本低。采用 SOSF 后，标准智能体开发中的文档预处理、结构化标注和元数据抽取环节工作量可大幅减少，有助于缩短开发周期、降低开发成本，具有较好的经济可行性。

3、预期达到的经济效果

1）降低标准数字化和智能体开发成本。采用 SOSF 后，标准文档可直接被 AI 智能体解析，减少人工结构化处理和重复录入，降低标准知识库建设成本。

2）推动标准知识普惠。SOSF 格式开放、易读、易解析，有助于中小企业、科研机构和社会公众便捷获取标准知识，缩小标准信息鸿沟。

3）减少纸质文档资源消耗。 标准开源格式推动标准以电子化、结构化方式发布和传播，有助于减少纸质标准印刷、分发和存储带来的纸张、油墨和物流能耗。

七、是否合规引用或者采用国际国外标准

该标准未采用国际国外标准。

八、与现行相关法律、法规、规章及相关标准的协调性

与现行相关法律法规、规章及现行有效标准无矛盾，协调统一。

九、重大分歧意见的处理经过和依据

无重大分歧意见。

十、涉及知识产权或专利的情况说明

无。

十一、实施国家标准的要求

建议将本标准作为推荐性团体标准。

九、贯彻标准的要求和措施建议

本标准的技术内容是推荐性。

建议本标准在发布后即开始实施。

十、废止现行相关标准的建议

无废止标准。

十一、其他应予说明的事项

无。