**团体标准《人工智能算力网络 第6部分：云际管理与调度》编制说明**

1. **工作简况**

2023年1月，根据中关村视听产业技术创新联盟2023年标准制修订计划的安排（AITISA【2023】1号文），下达了由鹏城实验室牵头制定团体标准《人工智能算力网络 第6部分：云际管理与调度》（标准计划号：2023011203）的制定任务。

本标准工作组自2022年3月第18次会议开始，讨论启动人工智能算力网络第6部分：云际管理与调度标准制定工作，通过了《人工智能算力网络 第6部分：云际管理与调度需求》提案，并输出需求文档。2022年6月AI标准工作组第19次会议通过提案征集书，输出《人工智能算力网络 第6部分：云际管理与调度 提案征集书》，会上经过讨论修改，通过了《人工智能算力网络 第6部分：云际管理与调度》草案提案，并输出《人工智能算力网络 第6部分：云际管理与调度 WD v0.1》。2022年12月第21次会议，会上经过讨论修改，通过了《人工智能算力网络 第6部分：云际管理与调度》草案提案，并输出《人工智能算力网络 第6部分：云际管理与调度 WD v0.2》。2023年12月第25次会议，会上经过讨论修改，通过了《人工智能算力网络 第6部分：云际管理与调度》草案提案，并输出《人工智能算力网络 第6部分：云际管理与调度 WD v0.3》。

2024年3月，经第26次会议讨论，同意团体标准《人工智能算力网络 第6部分：云际管理与调度》进入征求意见阶段，更新WD版本为0.4，输出征求意见稿。

本项目的主要起草单位包括鹏城实验室，北京大学，国防科技大学，华为技术有限公司，中科寒武纪科技股份有限公司。

本项目主要起草人：余跃、徐春香、杨晓珍、林冯军、王晖、张叶红、刘敏、曾炜、樊春、聂永丰、张鹏、朱宗卫、赵轩。

1. **标准编制原则和确定主要内容的论据及解决的主要问题**
2. **标准编制原则**

《人工智能算力网络 第6部分：云际管理与调度》需适用于人工智能算力网络的云际资源共享和联通调度相关的设计和建设，符合各行业人工智能应用对算力高效互联互通的需求趋势，体现人工智能算力网络的发展水平和方向，规定的技术要求结合产学研众多单位的意见，技术要求清晰、易操作、可实现。本部分的内容遵循T/AI 123.1—2023《人工智能算力网络 第1部分：总体要求》中对T/AI 123《人工智能算力网络》系列标准的规划。

**2. 标准主要内容**

本标准规定了人工智能算力网络（简称“智算网络”）的云际管理与调度的软件参考架构，规定了云际资源管理和作业调度的技术要求和调度策略要求，适用于智算网络中云际管理与调度相关功能的设计和开发。

本标准分为7个章节，第1章“范围”；第2章“规范性引用文件”；第3章“术语和定义”；第4章“缩略语”；第5章“软件参考架构”；第6章“功能要求”；第7章“接口要求”。

第1章阐述了本标准的主要内容及范围。

第2章说明了本标准的引用文件。

第3章列出了本标准中术语的引用和定义。从传统数据中心概念出发，进一步定义了“智算中心”、“智算网络”等专业术语。定义了云际管理与调度相关的“云际”、“资源”、“作业”等专业术语。

第4章列出了本标准涉及的缩略语。

第5章给出了智算网络调度层云际管理与调度系统的软件参考架构。

第6章提出了云际管理与调度系统的具体功能要求。

第7章提出了云际管理与调度系统的接口要求。。

**3. 解决的主要问题**

本标准拟解决以下问题：

1）不同区域的人工智能算力中心在系统架构、算力芯片、存储、网络资源上存在高度异构性，如何通过标准化手段屏蔽多层次异构性，将不同类型的泛在算力资源映射到统一的量纲维度，形成业务层可理解、可阅读的统一算力资源池，为智算网络的资源匹配调度提供基础保障成为亟待解决的技术和产业问题；

2）各智算中心内的数据要素建设相对分散，如何对数据资源进行存储、传输、权限的管理，如何与作业调度联动处理数据问题尚缺乏标准化指导；

3）各智算中心内的资源调度参考因素（简称“调度因子”）存在多分析维度、多数据来源和多数据规格等特性，如何统一这些调度因子的标准和规格，如何结合多维度的调度因子进行综合决策而达到最优的调度决策尚缺乏标准化指导。

1. **主要试验[或验证]情况分析**

本标准所规定的内容经过国内各大设备厂商及行业内实际研发过程中的试验和测试，已被证明确实可行。

1. **知识产权情况说明**

本标准不涉及任何专利及知识产权。

1. **采用国际标准和国外先进标准情况**

本部分未采用国际标准。

1. **与现行相关法律、法规、规章及相关标准的协调性**

本标准与现行相关法律、法规、规章及相关标准协调一致。

1. **重大分歧意见的处理经过和依据**

无。

1. **标准性质的建议**

建议发布为推荐性标准。

1. **贯彻标准的要求和措施建议**

本部分宜与T/AI 123《人工智能算力网络》系列标准搭配使用。

1. **替代或废止现行相关标准的建议**

无。

1. **其它应予说明的事项**

无。

团体标准《人工智能算力网络 第6部分：云际管理与调度》编制工作组

2024-03