

《标准机器语言表达 第2部分：基于XML的标准文档内容标记指南》（征求意见稿）

编制说明

一、工作简况

（一）任务来源

《标准机器语言表达 第2部分：基于XML的标准文档内容标记指南》是根据《国家标准化管理委员会关于下达2024年第二批推荐性国家标准计划及相关标准外文版计划》制定的，项目计划编号为“20240571-T-469”，计划完成时间为2025年。该计划项目由全国标准数字化标准化工作组（SAC/SWG 29）提出并归口。

（二）标准制定背景

2021年10月，中共中央，国务院印发《国家标准发展纲要》提出，要发展机器可读标准、开源标准，推动标准化工作向数字化、网络化、智能化转型。2021年12月，国家标准委、中央网信办、科技部等10部门联合印发《“十四五”推动高质量发展的国家标准体系建设规划》，提出深入推进国家标准数字化试点，探索增加机器可读标准、开源标准、数据库标准等新型国家标准供给形式；探索建立支撑国家标准数字化转型的信息系统。

当前，国内、国际标准化组织以及国内外标准化机构正在进行机器可读标准的研究。然而，数字化设备种类繁多，使用的机器语言并非一致，因此，目前尚无统一的面向机器可读、可执行需求的标准形式，导致标准实施仍然高度依赖人的理解，无法被机器直接使用，无法发挥数字化转型建设成果的最大价值；同时，标准的制修订者、使用者等不同用户处于不同的时空，在标准数字化生命周期不同阶段也很难及时共享成果。

通过制定本文件，可以为标准开发者、发布者、使用者提供一套通用的XML格式规范来发布和交换标准的内容及元数据，使其可以独立于标准内容的最初来源格式，从而可以为标准的发布和互用提供通用格式。进而能够产生如下效益：

从经济效益来看，可以节约标准数字化全寿期各个过程所需的时间，节省人力、物力成本。

从社会效益来看，能够使标准的制修订格式更加统一规范，同时，通过修改采用 IEC/ISO 的标准，也可以更好更快和国际接轨。

从生态效益来看，能够为我国标准数字化后续工作的开展奠定基础，促进我国标准数字化转型。

（三）起草过程

该项目下达后，由中国标准化研究院组织起草，于 2024 年 5 月广泛征集参编单位，组成涵盖产、学、研、用等相关单位的标准起草工作组，制定了标准编制工作计划，同步推进标准编制工作。

起草阶段工作如下：

2024 年 6 月 12 日，工作组在北京召开了起草工作组成立暨第一次工作会议，来自相关领域的科研院所、高校和企业 14 家单位的 20 余名代表参加会议。与会专家在对标准的适用范围及拟解决的问题、在标准体系中的定位和主要技术内容有了了解的基础上，明确了标准编制的基本思路与计划进度安排，就标准翻译的细节进行深入研究，并分为 5 个专项小组，分别进行讨论。

2024 年 7 月 5 日，工作组在北京召开了起草工作组第二次工作会议，与会专家参考 GB/T 1.2-2020，就采标过程中需要注意的事项进行了专题研讨，并就标准草案的翻译准确性、内容适用性及易用性等方面的问题进行了深入交流。标准草案经过工作组内部讨论、修订、完善，形成工作组讨论稿。

2024 年 8 月 14 日，工作组在北京召开了起草工作组第三次工作会议，会议通报了前期给工作组提交的材料中存在的共性问题，各组长介绍了标准草案主要修改内容及考虑因素，全体参会人员为标准草案修改过程中出现的问题进行了讨论。之后，各组根据会议结论，对工作组讨论稿进行了多次修改，形成了征求意见稿（初稿）。同时，工作组对形成的征求意见稿（初稿）在工作组内进行了征求意见。

2024 年 11 月 14 日，工作组在济南召开了起草工作组第四次工作会议，会上详细汇报了标准草案意见的收集情况，并对每条意见进行了逐一分析和讨论。工作组在充分听取各方意见的基础上，就标准草案意见处理达成了共识，最终形成了征求意见稿和编制说明。

在本标准起草阶段，SAC/SWG 29 工作组秘书处组织召开了内部会议 4 次，

就标准草案不同阶段的版本存在的问题进行讨论、记录，并及时将意见建议反馈给标准起草组，以便使本标准草案更加规范。

二、国家标准编制原则、主要内容和确定依据

（一）国家标准编制原则

标准编制遵循“一致性、协调性、适用性、实用性”的原则。本标准按照 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》规定的文字表述和基本格式编写，编制程序符合相关文件的规定，并与上述标准文件保持一致；

为了达到标准整体协调的目的，本标准在起草过程中参考了 GB/T 37967—2019《基于XML的国家标准结构化置标框架》、GB/T 42093.1-2022《标准文档结构化元模型 第1部分：全文》、GB/T 42093.2-2022《标准文档结构化元模型 第2部分：技术指标》等同类标准，并遵守 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》以及 GB/T 1.2-2020《标准化工作导则 第2部分：以ISO/IEC标准化文件为基础的标准化文件起草规则》的有关条款；

本标准的主体设计上考虑了采用ISO/IEC标准化文件的总体原则和要求，明晰了不同要素的编写方式，内容的表述便于标准使用者直接应用；

本标准的编制在符合ISO/IEC对于采标的相关要求的前提下，总结借鉴了不同行业、不同领域的实际需求，使标准制修订过程中使用XML的不同利益相关者能够有据可依，从而促进采标标准的规范化和应用效率。

（二）标准的主要内容

本文件提供了基于XML格式来标记标准化文件的文档结构、内容、表、数学公式、图、术语和定义的表示、引用等方面的指导和建议。

本文件适用于标准的编制者、制定机构和出版机构对国家、行业、地方、团体和企业标准进行基于XML格式的转换、编制、发布。

本文件正文共分为11章，其中前3章分别为范围、规范性引用文件、术语、定义和缩略语；第4章为本文件中涉及的标签、属性及ID方案，列出了本文件中出现的所有标签、属性以及本文件中主要的ID方案；第5章至第11章分别为文档结构、文档内容、表、公式、图、术语和定义的表示以及引用，分别阐述了

不同部分的标记方法。

本文件有 7 个资料性附录，1 个规范性附录，分别为：附录 A 章条编号对照一览表；附录 B 技术差异及其原因一览表；附录 C 本文件中涉及的标签及其定义的元素描述；附录 D 本文件中涉及标签的属性；附录 E ID 方案；附录 F 同类标准间标签名称的比较；附录 G 元数据的使用实例；附录 H 表的编码。所有附录均为资料性附录。

（三）标准主要内容的确定依据

本标准在编制过程中，修改采用了《IEC/ISO 编码指南》，吸收参考了《GB/T 37967-2019 基于 XML 的国家标准结构化置标框架》《GB/T 22373-2021 标准文献元数据》《GB/T 42093.1-2022 标准文档结构化 元模型 第 1 部分：全文》《GB/T 42093.2-2022 标准文档结构化 元模型 第 2 部分：技术指标》《标准机器语言表达 第 1 部分：基本架构与要素表达通用要求》等同类国家标准编制的规则和经验。

其中，根据我国国情及我国标准的特点，增加了第 3 章 术语、定义和缩略语，以及第 4 章 本文件中涉及的标签、属性及 ID 方案的内容；第 5 章 文档结构、第 6 章 文档内容、第 7 章 表、第 8 章 公式、第 9 章 图、第 10 章 术语和定义的表示和第 11 章 引用等章节基本按照原文内容进行编排。

同时，按本标准中出现的顺序对原文中附录的顺序行了编排，并根据采标的需要，增加了附录 A 章条编号变化对照一览表、附录 B 技术差异及其原因一览表、附录 C 本文件中涉及的标签及其定义的元素描述、附录 D 本文件中涉及的标签属性、附录 F 同类标准间元素名称的比较；删除了原标准中的附录 C 和附录 D；合并了原标准中的附录 E 和附录 F。

此外，本标准的采标原则如下：若原标准中只出现了关于 ISO 的使用方式及规则，则本标准中相关内容以 ISO 的方式呈现；若原标准中只出现了关于 IEC 的使用方式及规则，则本标准中相关内容以 IEC 的方式呈现；若原标准中同时出现了 ISO 和 IEC 的使用方式及规则，则本标准中相关内容以 ISO 的方式呈现。

三、与国际、国外同类标准水平的对比情况

本标准为修改采用 IEC/ISO 编码指南。

本标准水平为国内先进水平。

DITA (Darwin Information Typing Architecture) 标准是一种创建、组织和管理内容的 XML 标准,旨在提供一种通用的方法来定义和组织信息内容,使其更易于重用和共享。DITA 标准主要包括元素定义、语法规则和处理流程三个方面。与 DITA 标准相比,XML 是一种通用的标记语言,用于数据的传输和存储,具有广泛的适用性和灵活性;而 DITA 标准则是基于 XML 的用于技术文档的架构,具有面向主题的内容组织和高度模块化重用能力。

四、以国际标准为基础的起草情况,以及是否合规引用或采用国际国外标准,并说明未采用国际标准的原因

本标准修改采用 IEC/ISO 编码指南。

五、与有关的现行法律、法规和标准的关系

本标准遵循有关产业政策,符合国家相关法律、法规的规定,与现行法律法规没有冲突。

本标准隶属于全国标准数字化标准化工作组(SAC/SWG29),在标准体系中属于数字标准模型与定义—模型定义与开发—表达模型板块,本标准是数字标准的基本模型框架,是实现数字标准的基础。本标准提供了一种通用的 XML 格式来发布和交换标准的全文内容和元数据,适用于标准的编织者、制定机构和出版机构对国家、行业、地方、团体和企业标准进行基于 XML 格式的转换、编制和发布,使其具备可通过数字设备读取、传输、理解以及使用的问题,是实现标准数字化的关键技术手段。

本标准与 GB/T 37967—2019《基于 XML 的国家标准结构化置标框架》、GB/T 42093.1-2022《标准文档结构化 元模型 第 1 部分:全文》、GB/T 42093.2-2022《标准文档结构化 元模型 第 2 部分:技术指标》等同类标准,在侧重点上有所不同,本标准侧重于对标准文档内容结构化时所使用的标签进行标记。具体来说:

本标准提供了基于 XML 格式来标记标准化文件的文档结构、内容、表、公式、图、术语和定义的表达、引用等方面的指导和建议;适用于各类组织、机构对国家、行业、地方、团体和企业标准进行基于 XML 格式的转换、编制和发布。

GB/T 37967—2019《基于 XML 的国家标准结构化置标框架》规定了对国家标准文本进行结构化处理和表示时用到的命名空间前缀惯例、结构化处理基本架

构以及元素和标准定义的要求；适用于国家标准结构化的数字化加工以及相关开发与应用。

GB/T 42093.1-2022《标准文档结构化 元模型 第1部分：全文》从三个维度描述了标准文档结构化的全文元模型，即层次元模型、要素表述形式元模型、内容要素元模型，规定了各维度的关系及元模型元素的属性；适用于标准文档结构化信息化开发和应用。

GB/T 42093.2-2022《标准文档结构化 元模型 第2部分：技术指标》给出了标准中要求条款的性能指标描述的结构化元模型，规定了元模型的元素属性描述；适用于标准文档结构化开发与应用。

六、重大分歧意见的处理经过和依据

暂无。

七、涉及专利的有关说明

暂无。

八、实施国家标准的要求，以及组织措施、技术措施、过渡期和实施日期的建议等措施建议

建议本标准发布后3个月实施，通过标准宣贯会、研讨会等方式在标准制修订过程中使用XML的相关研制单位、用户单位、科研机构等进行广泛的宣传和推广应用。

九、其他应予说明的事项

根据《GB/T 1.2020 标准化工作导则 第2部分：以ISO/IEC标准化文件为基础的标准化文件的结构和起草规则》，本文件拟由等同采用修改为修改采用；同时根据我国实际情况，对原标准中不适用于我国的示例进行了删除；为了使翻译更加准确，拟将本标准的英文名称由Standard machine language expression - Part 2: Guidelines for standard document content labeling based on XML改成Standard machine language expression - Part 2: Guidelines for standard document content coding based XML。