ICS 03.220.40
CCS R 20

团　　体　　标　　准

T/CIN XXX—XXXX

绿色智能船舶协同研发平台
元数据服务管理规范要求

Requirements for metadata service management standard of green intelligent ship collaborative platform

（征求意见稿）

202X-XX-XX 发布　　202X-XX-XX 实施

中国航海学会　　发 布

目　　次

前　　言 1
1 范围 2
2 规范性引用文件 2
3 术语和定义 2
4 缩略语 3
5 基本原则 3
6 管理内容 3
7 元数据设计 4
8 元数据采集 4
9 元数据存储 5
10 元数据共享 5
11 元数据安全 6
12 元数据维护 6
参　考　文　献 8

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。

本文件由中国航海学会提出并归口。

本文件起草单位：上海船舶运输科学研究所有限公司、中远海运集装箱运输有限公司、中远海运能源运输股份有限公司、中远海运散货运输有限公司。

本文件主要起草人：屠海洋、乔继潘、黄珍平、张焱飞、陈伟民、季盛、李荣贵、陈超、张永亮、陈建挺、高玉玲、伍锐、文逸彦、顾红丽、冯亮、杜阳、李新颖等。

绿色智能船舶协同研发平台元数据服务管理规范要求

1 范围

本文件确立了绿色智能船舶协同研发平台元数据服务管理基本原则，规定了管理内容、元数据设计、采集、存储、共享、安全、维护等要求。

本文件适用于绿色智能船舶协同研发平台数据仓库设计与开发，同时适用于平台数据中心元数据管理及各类相关业务应用工作。

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 30522 科技平台 元数据标准化基本原则

3 术语和定义

下列术语和定义适用于本文件。

3.1

元数据 metadata

定义和描述其他数据的数据，可刻画数据属性，支撑数据存储定位、历史追溯、资源检索、文件记录等功能。

[来源：GB/T 18391.1-2009，3.2.16，有修改]

3.2

核心元数据 core metadata

描述数据最基本信息的元数据最小集合。

[来源：GB/T 30523-2023，3.3]

3.3

元数据元素 metadata element

元数据的基本单元，在单一元数据实体中具有唯一性。

[来源：GB/T 30523-2023，3.4]

3.4

元模型 meta model

规定一个或多个其他数据模型的数据模型。

[来源：GB/T 18391.1-2009，3.2.20]

3.5

业务元数据 business metadata

从业务角度描述数据资源各种属性的数据。

[来源：DB 1331/T 005-2022，3.3]

3.6

技术元数据 technical metadata

从技术角度描述数据资源各种属性的数据。

[来源：DB 1331/T 005-2022，3.4]

3.7

操作元数据 Operational Metadata
从操作角度描述数据处理和访问细节的元数据。
[来源：T/ZAII 035—2022，3.8，有改动]

3.8

管理元数据 Administrative Metadata
从管理协调角度描述数据资源各种属性的数据。
[来源：DB 1331/T 005-2022，3.5]

3.9

属性 attribute
一个对象或实体的特征。
[来源：GB/T 18391.1-2009，3.1.1]

3.10

绿色智能船舶协同研发平台 green intelligent ship collaborative platform
基于云计算、物联网、人工智能等数字技术构建的一体化协同系统（以下简称“平台”），整合船舶设计、制造、运维全链条资源，实现跨企业、跨地域的协同创新与数据互通。

4 缩略语

下列缩略语适用于本文件。
FTP：File Transfer Protocol（文件传输协议）
HDD：Hard Disk Drive（机械硬盘）
HTTP：HyperText Transfer Protocol（超文本传输协议）
MQTT：Message Queuing Telemetry Transport（消息队列遥测传输）
SSD：Solid State Drive（固态硬盘）

5 基本原则

元数据服务管理应遵循以下基本原则：
1） 一致性：建立统一的元数据标准、命名规范和口径规则，统一全域元数据的定义、逻辑与格式，确保表述、释义、规范统一；
2） 准确性：元数据需真实精准映射业务数据的属性、状态和特征，数据变更后及时校准，持续保障数据真实有效；
3） 完整性：全面覆盖数据业务、技术、操作、管理核心属性，完整归集必要要素，同时规避冗余、重复和内容冲突；
4） 可靠性：元数据来源可信、内容稳定、质量可控，体系具备良好的稳定性与可维护性，支撑平台长期业务应用与数据治理；
5） 可扩展性：管理架构与标准体系可适配数据量增长、类型迭代、数据源新增等业务变化，保障体系可持续适配迭代；
6） 可访问性：配置可视化、简易化的访问接口与操作工具，支撑合规用户高效完成元数据查询、浏览、调取等操作。

6 管理内容

元数据服务管理应至少包含以下内容：
1） 元数据设计：结合平台业务、技术、运行特性，规划搭建元数据体系，统一语义、字段属性与关联规则，确保定义规范、无歧义；

2） 元数据采集：采用自动化采集、人工校核补充相结合的方式，完成多源业务数据元数据的抓取、整理与归集；

3） 元数据审核：对新增、修订、变更的元数据，从一致性、准确性、完整性、可靠性四个方面开展合规校验审核；

4） 元数据存储：审核合格的元数据统一归档存储，建立规范化存储机制，保障数据安全稳定，支撑常态化业务调用；

5） 元数据共享：建立分级、可控、可追溯的共享体系，明确共享边界、权限与规范，实现合规前提下的资源有序复用；

6） 元数据安全：搭建全流程安全管控机制，覆盖存储、传输、访问、变更全环节，防范数据泄露、篡改、丢失风险；

7） 元数据维护：建立常态化巡检与动态维护机制，定期校准更新存量数据，变更数据严格执行审核流程，实现动态闭环管理。

7 元数据设计

7.1 设计原则

元数据设计贴合平台船舶研发业务实际，兼顾实用性与通用性，遵循以下原则：

1） 简单准确原则：采用行业通用术语，描述简洁规范、表意精准，规避冗余表述与晦涩技术话术；

2） 互操作原则：兼容国家、行业标准，适配多系统交互需求，保障数据互通、互认、互用；

3） 可扩展原则：架构灵活可迭代，可在不影响存量业务的前提下，新增、调整元数据元素与属性；

4） 需求导向原则：基于用户业务需求设计数据结构与语义规则，结合使用反馈持续优化，提升实用性。

7.2 设计步骤

元数据设计按照标准化闭环流程开展，具体步骤如下：

1） 分类框架规划：搭建业务、技术、操作、管理四类元数据分类体系，明确各类数据覆盖范围与资源边界；

2） 属性规范定义：对各类数据的业务、技术、操作、管理属性进行标准化定义，统一口径与释义；

3） 要素归集梳理：结合自动化工具与人工校核，全面归集元数据核心要素与技术参数，匹配业务真实状态；

4） 关联映射构建：梳理多模块、多系统数据关联关系，建立精准映射机制，形成完整元数据资源地图；

5） 动态迭代完善：结合平台业务迭代与场景更新，持续校核、补充、优化元数据，保障动态适配。

8 元数据采集

8.1 采集原则

元数据采集应遵循系统性、准确性、完整性、一致性和时效性原则，具体原则如下：

1） 系统性：明确采集范围，覆盖全部相关数据资源，无遗漏；

2） 准确性：真实反映数据资源属性，杜绝错误和失真；

3） 完整性：采集必备要素，补齐关键信息，支撑业务应用；

4） 一致性：同一数据资源在不同系统中的描述、格式和释义保持统一；

5） 时效性：建立常态化更新机制，及时同步变更信息，反映最新状态。

8.2 采集方法

结合业务场景与资源类型，可采用单一或组合方式开展采集：

1） 手动采集：人工编撰填报，适用于小众、特殊结构、无自动化采集条件的数据资源；

2） 自动提取采集：依托文本挖掘、语义解析等技术，批量自动生成标准化元数据，提升采集效率；

3） 标准化规范采集：依据国家、行业统一标准，规范元数据字段、格式与编码规则；

4） 数据库提取采集：对接业务数据库，自动读取表结构、字段、索引、关联关系等底层元数据；

5） 专用工具采集：使用专业元数据管理工具，实现数据批量录入、智能匹配与初步规整。

8.3 采集步骤

元数据采集应遵循规范化、全流程闭环管理，确保各环节衔接顺畅：

1） 需求分析与方案制定：明确采集目标、范围、质量要求及应用场景，制定采集方案，确定规则、方法与进度；

2） 资源筛选与队列建立：梳理存量与增量资源，确定采集清单，建立任务队列并明确优先级；

3） 元数据提取作业：按资源类型适配方法，有序提取核心及扩展元数据；

4） 元数据加工规整：开展清洗、转换、归一化处理，补齐缺失、修正错误、统一口径；

5） 元数据入库归档：将标准化元数据导入管理系统，完成关联匹配与索引建立，实现规范存储；

6） 动态更新与维护：建立常态化更新机制，随资源变更及时同步，全程记录变更信息，形成完整台账。

8.4 采集管理

全程规范采集全流程管控，确保采集质量与可追溯性：

1） 模型适配：采集规则、字段结构与平台元模型体系严格匹配；

2） 自动化优先：优先采用自动化采集模式，降低人工失误，提升标准化水平；

3） 任务调度：专业化管控采集时序、批次与优先级，支持第三方平台对接；

4） 日志追溯：完整记录采集任务执行状态、异常信息，实现全程可核查；

5） 结果推送：及时推送采集结果，保障运维与业务人员实时掌握工作进度；

6） 备份迭代：同步完成采集数据备份更新，规避数据不一致、丢失风险；

7） 版本管控：跟进元数据标准迭代，同步优化采集流程与校验规则。

9 元数据存储

9.1 基本要求

元数据采用集中统一存储模式，兼顾业务适配性与技术安全性，具体应符合下列要求：

1） 统一存储管控，解决数据分散、异构、格式杂乱问题，实现全域元数据集中管理；

2） 优先选用SSD高速存储介质，配备完善备份机制，保障存储稳定可用；

3） 存储系统支持弹性扩容，适配海量元数据存储与检索需求；

4） 采用结构化、层级化组织方式，建立多维度索引体系，支撑快速检索调用；

5） 配置数据加密、分级访问、定期备份等安全机制，保障存储安全；

6） 通过版本控制、资源锁控，规避多用户协同操作导致的数据冲突与篡改；

7） 全覆盖存储业务与技术元数据，完整保留各类数据属性信息；

8） 实行严格授权机制，严禁非授权人员查阅、篡改、删除元数据。

9.2 存储内容

元数据存储应包含以下内容：

1） 基础结构信息：数据名称、类型、格式、字段参数、约束规则等；

2） 业务定义信息：业务内涵、统计口径、计算逻辑、场景约束等；

3） 流转溯源信息：数据上游来源、下游应用场景与调用去向；

4） 关联关系信息：数据表、字段、数据集之间的映射、层级、联动关系；

5） 管理权限信息：数据归属、运维责任、更新频次、分级权限策略；

6） 合规约束信息：对应法律法规、行业规范、隐私保护等合规要求。

10 元数据共享

10.1 共享原则

元数据共享应遵循以下原则：

1） 按数据密级、类型、场景、用户权限实行差异化分级共享；

2） 立足用户实际需求提供共享服务，提升数据复用成效；

3) 严格遵守国家网络安全、行业数据规范，保障共享全流程合规；
4) 常规数据采用在线共享，高密级数据采用离线专属共享模式；
5) 公开类元数据需完成内部审批后方可对外发布；
6) 采用HTTP、FTP、MQTT等通用协议传输，保障数据完整安全。

10.2 共享要求

平台需具备完善的元数据共享服务能力，具体要求如下：
1) 制定并执行平台核心元数据规范，作为共享与管理的统一依据；
2) 所有对外共享元数据严格契合核心规范，保证标准化、统一化；
3) 设置便捷检索入口，支持用户快速查询各类数据资源；
4) 提供元数据在线浏览、配套下载、接口批量获取服务；
5) 平台各类数据库均需配备完整、规范的层级元数据信息；
6) 明确共享访问路径与使用规范，配套操作指引手册；
7) 建立共享标准迭代与问题反馈闭环机制，及时修正异常数据。

11 元数据安全

11.1 通用安全要求

元数据安全管理应满足以下通用要求：
1) 身份权限管控：严格身份核验，实行分级授权访问，杜绝未授权操作；
2) 敏感数据加密：对敏感元数据的传输、静态存储环节全程加密，防范信息泄露；
3) 完整性防护：部署校验机制，实时防范数据非法篡改、伪造；
4) 操作审计追溯：完整记录所有操作行为，留存关键日志，支撑安全核查；
5) 安全动态迭代：定期优化权限、加密策略等安全配置，适配风险防控需求；
6) 常态化备份：落实全量、增量备份机制，保障故障后快速恢复；
7) 安全培训：定期开展安全培训，提升从业人员风险防控能力。

11.2 备份专项要求

元数据备份应满足以下要求：
1) 定制专项备份计划，明确备份周期、范围、方式，常态化完成合规备份；
2) 采用异地多点冗余存储，规避单点故障导致的数据丢失；
3) 定期校验备份数据完整性、可用性，及时排查备份异常问题；
4) 严格管控备份文件权限，规范审批流程，严防未授权访问与泄露。

11.3 安全管控策略

应建立全流程元数据安全管控策略，确保元数据的有效管理和利用，应满足以下要求：
1) 搭建一体化安全防护体系，整合加密、备份、容灾、访问管控等防护手段；
2) 统一元数据标准化管理规范，实现全要素规范化管控；
3) 落实全生命周期安全管控，覆盖创建、存储、使用、迭代、废止全环节；
4) 常态化开展风险排查与隐患整改，筑牢数据安全防线。

12 元数据维护

12.1 基本维护要求

元数据维护应满足以下基本要求：
1) 全程记录维护操作轨迹，实现操作可追溯；
2) 建立问题反馈机制，及时处置元数据错误、缺失、不一致等问题；
3) 元数据变更必须执行审核审批流程，未经批准不得生效；
4) 规范版本管理，完整留存变更记录，实现版本可溯源；
5) 定期清理冗余、过时数据，降低运维复杂度；
6) 采用多节点副本部署，提升数据可用性与访问稳定性；

7） 常态化落实数据备份机制，防范数据安全风险。

12.2 一致性维护管控

为确保元数据的一致性，应采取但不限于以下管控措施：

1） 全局唯一标识管控，杜绝数据重复、冲突；
2） 独立版本号迭代管控，避免多节点更新数据错乱、覆盖；
3） 分布式协同管控，保障多节点数据同步一致；
4） 跨节点、跨系统同步校验，确保数据内容、格式、属性完全统一。

12.3 元数据校验

12.3.1 一致性校验

应校验元数据与源数据库原始信息的匹配一致性，校验内容包含：

1） 元数据核心内容与源数据一致；
2） 元数据字段数据类型与源数据定义一致；
3） 元数据描述信息与源数据业务定义一致。

12.3.2 完整性校验

应校验元数据信息的完整合规性，校验内容包含：

1） 元数据属性字段配置齐全、无缺失；
2） 规范要求的必填信息完整填报、无空缺；
3） 元数据整体内容完整合规、无遗漏。

12.4 元数据扩展

12.4.1 扩展原则

新建及优化元数据、开展元数据扩展工作时，应遵循以下基本原则：

1） 新建元数据的实体名称、元素名称及定义，不得与现有生效的元数据实体、元数据元素、代码表内容产生冲突、歧义与矛盾；

2） 制定专用元数据标准时，通用元数据标准中明确禁止扩展的内容，不得进行扩展；制定通用元数据标准时，核心元数据标准中明确禁止扩展的内容，不得进行扩展；

3） 新建元数据实体可组合复用现有元数据实体、元数据元素，同时可搭配新建元数据实体、元素构建；

4） 支持采用标准化代码表替换原有自由文本值域，优化规范已有元数据元素的值域定义；

5） 支持对现有代码表的元素进行新增、删减调整，扩充优化后的代码表需与原代码表保持业务逻辑一致性、延续性；

6） 支持对已有元数据的可选属性、最大出现次数等规则，设置更为严格的约束条件。

12.4.2 类型扩展

平台元数据支持以下合规扩展类型：

1） 属性扩展：新增元数据属性字段，丰富元数据维度信息；
2） 内容扩展：对元数据摘要、业务描述等文本内容进行补充、修订与完善；
3） 数据项扩展：新增元数据独立数据项，完善元数据体系。

12.4.3 扩展方法

元数据扩展流程、技术方法应按照 GB/T 30522 规定的元数据扩展规范执行。

12.5 元数据变更管理

12.5.1 业务数据调整后，需同步更新对应元数据，保障数据实时匹配；
12.5.2 支持单数据集精准修改、数据库批量更新两种变更模式，适配不同场景；
12.5.3 所有变更实行申报审批制，严禁直接篡改已发布元数据；
12.5.4 未通过审核的待发布数据可修改、撤回、删除，已发布数据仅管理员可删除；
12.5.5 开放专属应急变更功能，仅限核心管理员在应急场景合规使用。

参 考 文 献

[1] GB/T 18391-2009 信息技术元数据注册系统(MDR)
[2] GB/T 26499.3-2011 机械科学数据 第3部分：元数据
[3] GB/T 30523-2023 科技资源核心元数据
[4] GB/T 30535-2014 科技报告元数据规范
[5] GB/T 36478.3-2019 物联网信息交换和共享 第3部分：元数据
[6] GB/T 39316.2-2020 军民通用资源 元数据 第2部分：设备类 民用运输船舶
[7] CJJT 187-2012 建设电子档案元数据标准
[8] DB1310/T 279-2022 时空大数据平台服务元数据规范
[9] DB 1331/T 005-2022 雄安新区数据资源目录设计规范
[10] DB52/T 1540.2-2021 政府数据平台 元数据管理规范
[11] T/CASME 1246-2023 元数据采集管理技术规范
[12] T/PSC 7—2022 海洋科学数据共享服务规范 核心元数据
[13] T/ZAII 035-2022 数据中台 元数据规范