

S-97

S-100产品规范开发指南

1.1.0版-2020年6月

IHO



International
Hydrographic
Organization

由国际海道测量组织发布

中文版由中国海事局翻译



版权声明(S-97)

©版权所有国际海道测量组织所有(2020年)

本作品受版权保护,除《伯尔尼文学和艺术作品公约》(1886年)允许使用的用途及以下情况外,未经国际海道测量组织(IHO)事先书面许可,不得以任何方式翻译、复制、改编、传播或商业性利用本出版物的任何部分。本出版物中某些内容的版权可能归第三方所有,翻译和/或复制该内容须得到所有者许可。

作为一般参考,可在不超过成本的情况下,对本文件全部或部分进行翻译、复制或分发。未经 IHO 秘书处和任何其他版权所有者事先书面同意,不得以营利或收益为目的出售或分发副本。

如果根据上述条款对本文件全部或部分内容进行复制、翻译或分发,则应包括以下声明:

“IHO 秘书处代表国际海道测量组织(IHO)许可(许可证编号 10/23)复制 IHO 出版物《S-97 S-100 产品规范开发指南》(1.1.0 版),国际海道测量组织(IHO)不对复制材料的正确性承担责任。如有疑义,请以 IHO 的权威文本为准。含有 IHO 材料不得理解为 IHO 对本产品的认可。”

“本文件或出版物是 IHO 出版物 S-97 的译本。IHO 未对此译本进行检查,因此不对其准确性承担责任。如有疑义,请查阅 IHO 出版物《S-97 S-100 产品规范开发指南》。”

未经 IHO 秘书处事先书面许可,不得在任何衍生产品中使用 IHO 徽标(Logo)或其他标志。

文件历史

本规范由 IHO S-100 工作组协调修订。最新版本可以在 IHO 官网中下载。规范的维护应遵循 IHO 2/2007 决议及修正案。

版本号	日期	作者	备注
0.1	2018 年 1 月 31 日	EM, RM	第一版初稿
0.2	2018 年 8 月 31 日	EM, RM	根据 NIPWG 和 S-100WG 意见修订； 根据 S-100 4.0.0 版进行多处更新； 其他编辑性更新
1.0.0RC1	2019 年 3 月 11 日	RM, JW	IHO 秘书处编辑性修订和格式化修订； S-100 工作组第 4 次会议决定； 根据 S-100 4.0.0、S-121 和 S-123 参考更新
1.0.0	2019 年 5 月	S-100 工作组	第一版正式版
1.1.0	2020 年 6 月	S-100 工作组	文本修改,提高可读性； 第一版在 IHO 网站发布

目 录

A 部分 内容	1
A-1 概述	1
A-2 介绍	1
A-3 规范性引用文件	2
A-4 术语和缩略词	2
A-4.1 术语	2
A-4.2 缩略词	6
A-5 S-100 技术成熟度	7
A-6 S-100 产品规范模板及其构成	8
A-6.1 对产品规范可读性很重要的通用 S-100 概念	9
A-6.2 S-100 产品规范的主要组成	10
A-7 IHO S-100 基础设施	28
A-7.1 IHO 地理空间基础设施(GI)注册系统	29
A-7.2 要素目录构建器.....	29
A-7.3 DCEG 构建器	30
A-7.4 图示表达目录构建器.....	31
B 部分 编制产品规范的步骤	32
B-1 概述.....	32
B-2 介绍	32
B-3 规范性引用文件.....	32
B-4 术语和缩略词	33
B-4.1 术语	33
B-4.2 缩略词	37
B-5 开发步骤概述.....	37
B-5.1 基本开发过程.....	38
B-5.2 基于 S-100 的产品规范的审查周期	41
B-5.3 规范扩展过程.....	41
B-6 准备	42
B-7 开发数据模型(应用模式)	43
B-7.1 介绍.....	43
B-7.2 模型开发步骤.....	44
B-7.3 与通用要素模型的关系.....	46
B-7.4 应用模式规则.....	46
B-7.5 其他约定和建议.....	48

B-7.6	推荐做法	53
B-8	数据分类和编码指南	54
B-9	IHO GI 注册系统	55
B-10	要素目录	55
B-11	数据传输模式和封装	55
B-12	元数据	56
B-12.1	交换集产品元数据	56
B-12.2	服务元数据	63
B-13	定义数据编码格式	64
B-13.1	编码格式选择	64
B-13.2	GML 数据格式定义	65
B-14	图示表达元素和规则	65
B-15	图示表达元素的注册	66
B-16	图示表达目录	67
B-17	参考系统	67
B-18	数据产品交付	67
B-18.1	交付内容和结构	67
B-18.2	数据集更新	68
B-18.3	支持信息	69
B-19	有效性检查和数据质量	69
B-19.1	数据集的验证检查	70
B-19.2	包的有效性检查	70
B-19.3	通用的有效性检查	71
B-19.4	基本数据集与更新数据集的有效性检查	71
B-20	互操作性准备	71
B-21	样本数据/测试数据集	72
B-22	测试和反馈	72
B-23	工作流程	72
B-23.1	注册和请求 S 编号	72
B-23.2	项目组	72
B-23.3	作为开发过程的迭代优化	72
B-23.4	产品规范维护	73
C 部分	数据质量	74
C-1	概述	74
C-2	介绍	74
C-3	规范性引用文件	74
C-4	术语和缩略词	75
C-4.1	术语	75
C-4.2	缩略词	76
C-5	数据质量度量概述	77

C-5.1 数据质量度量.....	78
C-6 给产品规范开发人员的建议	79
C-6.1 完整性 > 冗余/缺失	79
C-6.2 逻辑一致性 > 概念一致性	79
C-6.3 逻辑一致性 > 域一致性	80
C-6.4 逻辑一致性 > 格式一致性	80
C-6.5 逻辑一致性 > 拓扑一致性	80
C-6.6 位置精度	80
C-6.7 主题准确性	82
C-6.8 时间质量	82
C-6.9 聚合	82
C-6.10 可用性	83
C-7 数据质量度量.....	84
C-8 数据验证的最低标准	88

A 部分 内容

A-1 概述

S-100,即通用海道测量数据模型,是一种支持各种海道测量相关的数字数据源的海道测量地理空间数据标准,且完全符合主流的国际地理空间标准,特别是 ISO 19000 系列地理信息标准。这种对标使海道测量数据和应用程序更容易集成到地理空间解决方案中。S-100 相比 S-57 更为灵活,为使用栅格和网格数据类型、增强型元数据和多种编码格式提供了条件。它还通过专门的在线注册表为要素、属性和图示表达提供了更灵活和动态的维护机制。S-100 提供了一个为海道测量数据建立标准化产品规范的组件框架,从而为不同的数据标准和系统之间提供真正的互操作性。

本指南旨在为所有计划开发或扩展符合 S-100 标准的产品的人员提供指导。

A-2 介绍

创建基于 S-100 的产品规范对于缺乏 S-100 研究经验的团队来说可能是一个很大的挑战,主要原因在于 S-100 是一个涵盖广泛应用的框架标准,它可能不会在每个产品规范中使用。S-97 由 IHO 创建,旨在帮助产品规范开发人员更好地理解 S-100,并为创建和扩展基于 S-100 的产品规范提供指导。

本指南的核心目标之一是协助创建可用于 e-航海生态系统的统一产品规范。e-航海生态系统的含义是指为 IMO 定义的船岸 e-航海系统(如电子海图显示与信息系统,ECDIS)而制定的所有产品规范,但 S-100 可以扩展到 e-航海之外,有更广泛的应用范围。

本指南旨在为所有计划开发或扩展符合 S-100 标准的产品的人员提供指导。本标准由以下三部分组成:

- A 部分(本文档)是对基于 S-100 的产品规范的各个组件的深入描述;
- B 部分(单独的文档)描述了创建基于 S-100 的产品规范所涉及的典型步骤和活动。B 部分描述了整个过程、具体活动和任务,包括对产品规范开发过程中需要解决的具体问题的提示;
- C 部分(单独的文档)描述了适用于基于 S-100 的产品规范的数据质量度量。

A-3 规范性引用文件

IHOIB IHO S-100 信息手册,2017年5月。

ISO 8211 ISO 8211 信息交换结构用数据描述文件规范,ISO/IEC8211,1994年。

ISO 646 信息技术 国际编码标准信息交换用 ISO 7 位编码字符集,ISO/IEC646,1991年。

ISO 10646 信息技术 通用编码字符集(UCS),ISO/IEC10646,2017年。

ISO 19103 地理信息 概念模式语言,ISO 19103,2005年。

ISO 19110 地理信息 要素编目方法,ISO 19110,2005年。

ISO 19115-1 地理信息 元数据 第1部分:基础,ISO 19115-1,2014年,2018年第一次修订版。

ISO 19136 地理信息 地理标记语言(GML),ISO 19136,2007年。(也可作为 OGC 07-036 地理标记语言(GML)编码标准获得。开放地理空间联盟公司,2007年。)

S-57 IHO 数字海道测量数据传输标准 3.1 版,2000年11月。

S-58 IHO S-58 电子海图有效性检查 6.0.0 版,2017年5月。

S-99 IHO S-99 S-100 地理空间信息注册系统组织与管理操作程序 1.1.0 版,2012年11月。

S-100 IHO S-100 通用海道测量数据模型 4.0.0 版,2018年12月。

S-122 IHO S-122 海洋保护区 1.0.0 版,2019年1月。

S-123 IHO S-123 海上无线电服务 1.0.0 版,2019年1月。

注:除非特别标明版本,则本文档提及的 S-100 均指 S-100 4.0.0 版。

A-4 术语和缩略词

A-4.1 术语

抽象类 **abstract class**

无法实例化的对象类,或在信息模型中被指定为不允许实例化的对象类。

注:抽象类的子类可以是抽象的,也可以是非抽象的。

聚合 **aggregation**

一种特殊的关联形式,用于指定聚合体(整体)和组成部分(见组合)之间的整体与部分关系。[ISO 19103]

应用 **application**

支持用户要求的数据操作和处理。[ISO 19101-1:2014]

应用模式 application schema

一个或多个应用所需数据的概念架构。[ISO 19101-1:2014]

关联 association

两个或多个类之间的语义关系,用于指定它们的实例之间的连接关系。[ISO 19103]

属性 attribute

(1)实体的命名特性。[ISO/IEC 2382 - 17:1999]

注:描述实体的几何、拓扑、主题或其他特征。

(2)UML:分类器中要素实例可能包含的值的范围。[ISO/IEC 19501:2005(改编)]

(3)XML:元素中包含的名称与值对。[ISO 19136]

base64 编码

一种用于表示任意序列的二进制八位数编码,其格式允许同时使用大小写字母,但不必是人类可读的。

[IETF RFC 4648(重新设计)]

代码 code

遵循特定方案的表示标签。[ISO 19118:2011]

代码表 codelist

包含允许值代码的值域。[ISO 19136]

代码空间 codespace

代码、名称、术语或类别的规则或权限。[ISO 19136]

示例:codespace 的示例包括字典、权威机构、代码表等。

组合 composition

作为整体中的一部分,具有强所有权和相同生命周期的聚合关联形式。[ISO 19103]

概念模型 conceptual model

定义话语世界概念的模型。[19101-1:2014]

概念模式 conceptual schema

概念模式的正式描述。[ISO 19101-1:2014]

数据客户端 data client

接收基于 S-100 的加密数据的终端用户。数据客户端将使用软件应用程序(例如 ECDIS)来执行 S-100 保护方案中详细描述等诸多操作。[S-100 第 15 部分(改编)]

例如:ECDIS 用户。

数据许可证 data permit

包含解密许可产品所需的加密产品密钥的文件,通常是专门为特定数据客户端创建的。[S-100 第 15 部分(改编)]

数据服务器 data server

生产加密数据文件或向数据客户端颁发数据许可证的组织。[S-100 第 15 部分(改编)]

要素 feature

真实世界现象的抽象表现。[ISO19101:2003]

注:要素可能以类型、类或实例的形式出现。要素类型或要素实例仅用于一种情况时应予以明确。要素类应在模型或应用程序架构的语境中使用。

示例:被称为“埃菲尔铁塔”的真实世界现象可以与其他现象一起归入名为“铁塔”的要素类型。

要素关联 feature association

将一种要素类型的实例与相同或不同要素类型的实例链接的关系。[ISO 19110]

要素属性 feature attribute

要素的特征。[ISO 19101]

注:一个要素属性类型有一个名称、一个数据类型和一个与其相关联的域。要素属性实例具有从要素属性类型的值域获取的属性值。

示例 1:名为“颜色”的要素属性可能具有属于数据类型“文本”的属性值“绿色”。

示例 2:名为“长度”的要素属性可能具有属于数据类型“实数”的属性值“82.4”。

要素目录 feature catalogue

包含定义和描述一组或多组地理数据中出现的要素类型、要素属性和要素关联的目录。[ISO 19110]

要素类 feature class

应用模式或模型中表示要素的类。

标识符 identifier

一种语言上独立的字符序列,能够唯一和永久地识别与之相关的字符。[改编自 ISO/IEC 11179-3:2003]

信息类型 information type

仅具有主题属性的数据集中可识别的信息单元。[S-100 3-5.1.2(改编)]

示例:信息类型可以包含海图改正信息。

注:信息类型也可以相互关联。在有与信息类型有关的补充信息或需要翻译信息时,可以这样处理。例如,携带海图改正的信息对象可以包含英文文本,相关联的补充信息对象可以携带相应的中文文本。

实例化 instantiate

通过实例表达。[韦氏词典在线版 <https://www.merriam-webster.com/dictionary/instantiate>]

接口 interface

表征实体行为的命名操作集。[ISO 19119:2005]

元数据 metadata

有关资源的信息,[ISO 19115-1];定义和描述其他数据的数据。[ISO 11179-3:2013]

模型 model

现实世界某些方面的抽象。[ISO 19109-2015]

操作 operation

可以调用对象来执行的转换或查询的规范。[ISO 19119:2005]

注:一个操作有一个名称和一个参数列表。

注册表 register

一组文件,其中包含分配给具有相关项描述的项的标识符。[ISO 19135]

注:描述可能包括许多类型的信息,包括名称、定义和代码。

注册系统 registry

维护注册表的信息系统。[ISO 19135]

关系 relationship

模型元素之间的语义联系。[ISO 19103]

资源 resource

符合要求的可识别资产或手段。[ISO 19115-1]

示例:数据集、数据集系列、服务、文档、计划、软件、人员或组织。

方案管理员 scheme administrator

专门负责维护和协调 S-100 规定的保护方案的组织。[S-100 第 15 部分(改编)]

服务 service

实体通过接口提供的功能的不同部分。[ISO 19119:2005]

空间对象 spatial object

用于表示要素的空间特性的对象。[ISO 19107:2003]

流 stream

在线数据交换中,由通信系统传送的连续的碎片化数据序列。[S-100]

话语世界 universe of discourse

对包括所有感兴趣事物的现实或假想世界的看法。[19101-1:2014]

词汇 vocabulary

包含一个或多个特定主题领域的名称和定义的术语词典。[ISO 1087-1:2000]

A-4.2 缩略词

- AIS 自动标识系统(Automatic Identification System)
- DQWG 数据质量工作组(Data Quality Working Group)
- ECDIS 电子海图显示与信息系统(Electronic Chart Display and Information System)
- ENC 电子海图(Electronic Navigational Chart)
- GML 地理标记语言(Geography Markup Language)
- GFM 通用要素模型(General Feature Model)
- GI registry 地理空间信息注册系统(Geospatial Information registry)
- HDF 层次数据格式(Hierarchical Data Format)
- HTTP 超文本传输协议(Hypertext Transfer Protocol)
- HTTPS HTTP 安全(HTTP Secure)
- IALA 国际航标协会(International Association of Lighthouse Authorities)
- IEC 国际电工委员会(International Electrotechnical Commission)
- IHO 国际海道测量组织(International Hydrographic Organization)
- IMO 国际海事组织(International Maritime Organization)
- ISO 国际标准化组织(International Organization for Standardization)
- RENC 区域电子海图协调中心(Regional ENC Coordinating Centre)
- REST 表述性状态转移(Representational State Transfer)
- SENC 系统电子海图(System Electronic Navigational Chart)
- SOAP 简单对象访问协议(Simple Object Access Protocol)
- SOS 传感器观察服务(Sensor Observation Service)
- SSL 安全链路层(Secure Sockets Layer)
- SVG 可缩放矢量图形(Scalable Vector Graphics)
- TCP/IP 传输控制协议/因特网协议(Transmission Control Protocol/Internet Protocol)
- VTs 船舶交通服务(Vessel Traffic Service)
- WSDL 网络服务描述语言(Web Services Description Language)
- WFS 网络要素服务(Web Feature Service)
- XML 可扩展标记语言(eXtensible Markup Language)
- XSD XML 模式定义(XML Schema Definition)
- XSLT 可扩展样式表语言转换(eXtensible Stylesheet Language Transformations)

A-5 S-100 技术成熟度

在 S-100 框架下开发新产品规范的关键在于与更广泛的社会组织或团体沟通规范的完备性及其可操作性。由于正在开发的产品规范具有多种不同类型的操作设置,并且不是所有规范都需要 S-100 的所有组件,这也使其更加复杂。因此,海道测量服务和标准委员会(HSSC)采用了 S-100 技术成熟度的概念。

技术成熟度的概念展示了一个从想法到常规应用的过程,它可以让 IHO 成员清楚地了解该规范是否可以得到认可和批准。也可以让其他使用 S-100 框架的非 IHO 利益相关组织衡量其产品规范何时才能符合从过渡到实际操作的技术成熟度。

表 A-5-1 列出了符合每个 S-100 技术成熟度的先决条件。需要注意的是,所有基于 S-100 的产品规范都必须符合 S-100 标准,并且必须使用已发布的 S-100 基础设施和流程对要素目录和图示表达目录进行创建和维护。

S-100 技术成熟等级

表 A-5-1

所需产品规范组件	1 级 v1.0.0	2 级 v1 - 2.0.0	3 级 >v2.0.0	4 级 >v2.0.0	5 级 >v2.0.0
主文档(定义了产品规范要求的 S-100 中相关部分)	×	×	×	×	×
默认编码	×	×	×	×	×
S-100 兼容要素目录	×	×	×	×	×
	(草案)	(更新)	(定稿,通过 IHO GI 注册)		
数据分类和编码指南	×	×	×	×	×
	(草案)		(定稿)		
S-100 兼容描述目录					
注:并非每个规范都需要描述目录,这应该作为开发过程和涉众反馈的一部分来确定		×	×	×	×
数据质量检查		×	×	×	×
测试数据集		×	×	×	×
数据验证(及测试数据集)		×	×	×	×
交换目录		×	×	×	×
加密/数字签名			×	×	×
互操作性			×	×	×
			(草案)	(经测试)	*

续上表

所需产品规范组件	1级 v1.0.0	2级 v1-2.0.0	3级 >v2.0.0	4级 >v2.0.0	5级 >v2.0.0
警告和显示				x*	x*
操作数据					x

(x*:仅用于 ECDIS)

1级:包含开始测试数据集和系统原型开发所需的最少数量的组件,1级是演示开始前的最后一个开发阶段,通常是产品规范的1.0.0版。

2级:包括数据质量检查和测试数据集等附加项,以便产品规范可以在原型环境中演示。通常以产品规范版本1.n.n~2.0.0的版本形式出现。根据产品规范的最终用户要求,2级可以在操作语境中实现。随后的S-100技术成熟度取决于导航系统内产品的操作要求。

3级:构建在2级之上,包括一个功能齐全且文档化的交换目录,以及(可选的)用于数据和实现系统的加密层。在这个级别,应在真实的环境中演示原型系统、产品或流程。

4级:仅用于船舶导航系统,如ECS和ECDIS。在本级别,产品规范的开发人员需要确保根据S-98的要求对互操作性、报警和提示功能进行思考并形成文件。在本级别,应该有一个基础、合规的系统、过程或产品以显示其预期的运行或功能。

5级:系统、过程或产品已部署和使用。在本级别,数据和合规系统随时可供操作使用。此阶段包括基于机器的现势性检查功能(即自动预警和更新状态报告),现势性所需的功能可以在个别产品规范中提供,也可以通过S-128航海产品目录提供。

A-6 S-100 产品规范模板及其构成

数据产品规范是定义地理空间数据产品的精确技术描述。它描述了给定应用程序的所有要素、属性和关系,以及它们到数据集的映射。包括用于数据识别的一般信息以及以下信息:

- 数据内容和结构;
- 参考系统;
- 数据质量方面;
- 数据收集;
- 图示表达;
- 维护;
- 交付;以及

- 元数据。

S-100 第 11 部分描述了地理数据产品的数据产品规范。它的目的是为将要编写的任何数据产品规范提供一个清晰和相似的结构。产品规范应构成一组人类可读的文档。通常,它还应包括机器可读文件,以获取要素目录、图示表达目录等信息。符合产品规范的示例见 S-100 的附录 11-B。

A-6.1 对产品规范可读性很重要的通用 S-100 概念

A-6.1.1 强制性要求与可选要求

为了符合 S-100 的要求,产品规范必须包含一些特定的组成部分。例如,它需要包括一个要素目录,而图示表达目录是可选的。S-100 利用多重字段来帮助确定产品规范中哪些是强制性的。表 A-6-1 就是其中一个例子。

S-100 表示示例(数据产品信息描述,S-100 表 11-1)

表 A-6-1

名称	描述	多重性	类型
标题	数据产品的官方名称	1	字符串
摘要	数据产品的非正式信息描述	1	字符串
缩写	数据产品标题的简称	0..*	字符串
内容	遵循规范的数据集的内容描述	1	字符串
空间范围	数据产品覆盖的空间范围描述	1	EX_Extent (ISO 19115-1)
时间范围	数据产品覆盖的时间范围描述	0..1	EX_Extent (ISO 19115-1)
特定用途	数据的用途	1	字符串

在上面的示例中,多重性一栏表示哪些元素是必需的,哪些是可选的。S-100 使用以下符号来表示多重类型:

- 1 必选
- 1..* 至少一个或多个
- 0..1 可选或只有一个
- 0..* 可选或多选

A-6.1.2 驼峰命名法及其在 S-100 中的使用

S-100 广泛使用驼峰命名法(camelCase),其使用基于 ISO 19103 原则。驼峰命名法(通常写为 camel-Case 或 CamelCase)是一种书写复合词或短语的做法,即短语中间的每个词或缩略语以大写字母开头,中间没有空格或标点符号。常见的例子包括“iPhone”“eBay”“FedEx”“DreamWorks”“HarperCollins”“iCarly”“WordWorld”“WordGirl”等。S-100 使用驼峰命名法作为构造 S-100 本身和 IHO 地理信息(GI)注册系统中使用的元素的不同标识符或名称的一种方法。

驼峰命名法标识符必须符合 S-100 条款 2a-4.2.3 的如下要求：

- 是一种复合词,单词之间没有空格相连,并且复合词中的首字母大写;
- 在注册表中是唯一的;以及
- 符合 ISO 10646¹,大写字符为 A~Z,0~9,“_”和小写字符为 a~z;

此外:

- 要素和信息类型必须以大写字符 A~Z 开头;
- 复杂和简单属性、代码表、属性和枚举值必须以小写 a~z 开头。

示例 1:BeaconCardinal 是要素 Beacon Cardinal 的驼峰命名法标识符。

示例 2:categoryOfLandmark 是陆标类的驼峰命名法标识符。

A-6.2 S-100 产品规范的主要组成

本节重点介绍构成 S-100 产品规范的组件,并阐述为什么需要这些组件。

A-6.2.1 概述部分及其分项内容

产品规范的概述部分为读者提供有关数据产品的一般介绍性信息以及产品规范元数据。S-100 指出,概述应包括以下部分(不包括“引用文件”和“语言的使用”,但本指南建议也包括这两部分):

- 介绍;
- 引用文件;
- 术语、定义和缩略词;
 - 语言运用
 - 术语和定义
 - 缩略词
- 数据产品描述概要;
- 数据产品规范元数据;
- 产品规范维护。

A-6.2.1.1 介绍

本节提供有关创建产品规范的信息,其中包括产品规范的主题和意图。

A-6.2.1.2 引用文件

本节应列出定义产品规范的整体要素、实现产品规范所依据的标准,如规范性的 ISO 或其他标准。包含其他有用信息的标准或文件,如果不是产品规范的组成部分,可以作为参考文件。

¹ S-100 2a-4.2.3 指定 ISO/IEC 646(ASCII),而其他地方使用 10646(UTF-8)。由于 UTF-8 和 ASCII 混用可能存在问题,本指南建议只使用 UTF-8。

A-6.2.1.3 语言运用

语言运用虽然是个可选项,但添加一个语言运用部分来详细说明产品规范文档(包括附录和附件)中使用的特定词语的预期含义是有益的。其目的是尽可能消除这些词的模糊性,以便在规范中清楚地说明什么是强制性要求,什么是强烈推荐的以及什么是可选的。IHO 的规范中经常使用以下语句:

- Must 表示强制性要求。
- Should 表示可选要求,即建议遵循的流程,但不是强制性的。
- May 是指“允许”或“可能”,不是强制性的。

A-6.2.1.4 术语和定义

术语和定义是有用的参考,反映了规范的内容及其预期用途。

A-6.2.1.5 缩略词

规范中使用的缩略词都应在产品规范介绍的缩略词部分中列出其全部含义。

A-6.2.1.5.1 首字母缩写

按照惯例,数据产品的名称要用首字母缩写,例如 AML(Additional Military Layer,附加军事层)或 ENC(Electronic Navigational Chart,电子海图)。由于各种原因,缩略词也可以在整个规范中使用,这些缩略词应该编写在文档开头的缩略词段落中,作为读者的快速参考使用。

A-6.2.1.6 数据产品描述概要

这部分是对数据产品的非正式描述,也可以作为规范的摘要使用,用以明确产品的预期使用目的。另见第 A-6.2.4。

A-6.2.1.7 产品规范维护

IHO 发布的产品规范变更将作为新版本、修订版或包含澄清说明的文档发布。使用哪个级别取决于更改的性质。其他发行人很可能会效仿 IHO 的做法。一般而言,本文遵循 S-100 第 12 部分中的指导,其中描述了三种类型的变化,即新版、修订和澄清。这些变更类型如下所述。

A-6.2.1.7.1 新版

新版引入了重大变化。新版本允许引入新概念,例如支持新功能或新的应用程序,或引入新的构造或数据类型。新版对于产品规范的当前用户和潜在用户均会产生重要影响。

A-6.2.1.7.2 修订

修订版被定义为实质性的语义变化。通常情况下,修订将引入更改以纠正相关错误,或引入由于实践经验或环境变化而导致的必要更改。修订不是澄清,修订会对当前用户和潜在用户产生影响。修订版在发布时必须包含此前累积的所有澄清说明。

修订版中的更改可确保同一版本中的新版本与旧版本兼容。例如,较新的修订引入了新的要素和属性。在同一版本中,一个版本的数据集始终可以与要素和图示表达目录的较新版本一起处理。在大多数情况下,一个新的要素或图示表达目录将导致该规范的修订。

A-6.2.1.7.3 澄清

澄清是非实质性改动。通常情况下,澄清可以消除歧义,纠正语法和拼写错误,修改或更新交叉引用,插入改进的图形、拼写、标点符号和语法等。澄清不得引起任何实质性的语义变化。

澄清中的更改很小,并确保与同一版本中的新版本与旧版本兼容。在同一版本中,可以始终使用更高版本的要素和图示表达目录处理前一个版本的数据集。而图示表达目录始终可以依赖于要素目录的较早版本。

A-6.2.2 版本号

用于标识 S-100 和派生规范中的更改(n)的相关版本控制编号通常遵循以下格式:

- 新版表示为 n.0.0;
- 修订版表示为 n.n.0;
- 澄清版表示为 n.n.n。

IHO 其他标准也多采用了同样的版本控制格式。

A-6.2.3 规范范围

规范范围是指产品规范的指定部分对整个产品或部分产品的适用性。

例如,坐标参考系统通常适用于整个产品,而导航要素和语境特征的维护机制可能不同。这种差异可以用规范范围进行描述。根据产品规范的类型,范围可以包括表 A-6-2 中的项目。

规范范围信息(S-100 表 11-3)

表 A-6-2

名称	描述	多重性	类型
范围标识 scopeIdentification	范围的具体标识	1	字符串
级别 Level	范围规定的数据的层次级别	0..1	MD_ScopeCode (ISO 19115-1)
级别名称 levelName	规范规定的数据的层次级别的名称	0..1	字符串

续上表

名称	描述	多重性	类型
级别描述 levelDescription	级别的详细说明	0..1	字符串
覆盖范围 Coverage	要素的子类型,使用一组属性来表达真实世界现象	0..1	字符串
范围 Extent	数据的空间、垂直和时间范围	0..1	EX_Extent (ISO 19115-1)

如果一个规范在整个数据产品中是同质的,那么只需要定义一个通用范围(根范围),产品规范的每个部分都适用于此通用范围。此通用范围可能类似于以下示例:

范围标识:全局范围

级别:006-series

级别名称:数据集

级别描述:级别适用于所有 ENC 数据集

覆盖范围:ENC 要素目录中的所有要素

范围:EX_GeographicBoundingBox

westBoundLongitude: - 180

westBoundLongitude: 180

southBoundLatitude: - 90

northBoundLatitude: 90

级别属性是 ISO 19115-1 中的一个代码列表,称为 MD_ScopeCode,包含了规范的主要组件。范围属性是一个类,可以是描述性字符串、地理范围(如上例所示)、垂直范围或时间范围的任意组合。

产品规范可以基于一个或多个标准来指定产品的数据内容的分区。对于产品规范的不同部分,这种分区可能是不同的。数据内容的每一部分都应由一个规范范围来描述,该规范范围可以继承或覆盖通用范围规范。原则上,产品规范的任何或所有剩余部分可以具有适用于产品内范围的变体。每个变体必须标识其适用的范围。

示例:支持导航的数据产品通常包含两组要素类型,即提供快速变化的导航信息的要素类型和对导航安全至关重要的要素类型,以及提供背景参考信息的项。维护和交付信息将根据这些分组进行划分,对必要信息将予以维护和交付,而对参考系统信息则不予以维护和交付。

A-6.2.4 数据集标识

除了对数据产品的非正式描述(另见第 A-6.2.1.6 节)外,S-100 还要求有一节用以描述根据特定产品规范系列创建的能够唯一标识任何数据集的信息。

与数据产品的概述信息不同,数据集标识用于单个数据集。如有必要,可以对其中的一些元素进行标准化。例如,用途属性值可能通用在根据特定规范创建的所有数据集。其他属性可能受益于遵循通用模式,例如数据集标题可能遵循用户熟悉的特定样式。一些属性是定义在其他地方的代码列表类型,如ISO19115-1。这些属性被限定为代码列表中给出的值,例如,并非所有值对于规范的范围都有意义,那么在某些情况下,将给定的代码表限制为值的子集是有益的。

这些信息存储在与数据集关联的元数据中。因此,必须确保提供适当的元数据属性,从而使本节与元数据节协调一致。

一些产品规范已经将数据产品的非正式描述与数据集标识部分合并到一个部分中,这是被允许的可选项。

A-6.2.5 数据内容和结构(应用模式)

应用模式定义了S-100产品的数据内容和结构。应用模式使用S-100第1部分(概念模式语言)中描述的统一建模语言(UML)来表达,并允许开发人员以一致和可维护的方式实现S-100产品规范。S-100的通用要素模型(第3部分)规定了开发应用程序模式的规则,其中包括要素的概念模型及其特征和关联。

S-100产品规范模板规定,具有大型应用程序模式的规范只需在规范文档中包含UML中的特定示例,因为应用程序模式需要在要素目录中完整的实现。由于应用程序模式通常变得太大,无法在一个页面中保持可读性,因此根据函数和元素,将应用程序模式拆分为多个部分可能更有利于整体的可读性。

A-6.2.5.1 基于要素的数据内容结构

地理应用的数据内容是根据真实世界要素的视角并结合特定应用程序的需求定义的。内容按照对象进行构建。S-100考虑了第3部分第5.1款中定义的两类类型的对象或要素,定义如下:

- 1)要素:要素与其属性一起定义。要素是对现实世界现象的抽象表现。要素包含两个方面:要素类型和要素实例。要素类型是一个类,在要素目录中定义。要素实例是要素类型的单个实例,并表示为数据集中的对象。
- 2)信息:信息类型用于在要素和其他信息类型之间共享信息。信息类型是在要素目录中定义的一类对象。信息类型的实例是数据集中可识别的信息单元。信息类型仅具有主题属性。信息类型的实例可以与一个或多个要素实例或其他信息类型的一个或多个实例相关联。信息类型可以被视为共享属性。

通用要素模型(GFM)为这些对象提供了一个概念模型。要素类型的定义保存在要素目录中。GFM 还充当要素目录的概念模型。空间信息在 S-100 第 7 部分空间模式中定义,由简单的几何图形组成,这些几何图形可以用多种配置表示。应用模式必须定义产品规范中使用的空间组件以及与要素类的关系。

A-6.2.5.2 基于覆盖的数据内容结构

S-100 还将图像、网格和覆盖数据定义为面向要素的数据。在最简单的形式中,可以将图像或任何一组网格化数据视为单个要素。因此,用于要素数据的应用模式的规则也适用于图像和网格化数据。但是,必须注意确保应用模式能够根据 S-100 第 8 部分第 8-6 条的规定以及第 8-8 条中定义的网格化数据空间引用要求来准确地定义图像和网格。如果产品包含一系列或一组图像或网格化数据集,则应按照 S-100 第 8 部分第 8-7 条中定义空间关系的应用模式来规定定义。

A-6.2.6 数据产品格式

基于 S-100 的产品规范应定义数据产品中每个范围的交付格式(编码)。这包括对文件结构和格式的描述(如果适用)或数据流的格式。编码结构可以在产品规范指定,或者通过参考单独的专用标准配置文件或标准来指定。S-100 包括三种编码的专用标准配置文件,即 ISO 8211 二进制编码、GML(ISO 19136)编码和 HDF5 编码。产品规范可以引用这些专用标准配置文件,并说明如何在特定的产品规范中使用这些标准。例如,给定的产品具有一个特定的 GML 应用模式,并使用一个或多个 XML 模式定义语言文件来表示。

专用产品可以通过在产品规范中指定整个编码(或通过引用已建立的外部标准,或两者的适当组合)来使用其他编码。在这种情况下,实施成本可能高于使用标准 S-100 编码的系统。

S-100 专用标准配置文件编码的简要说明见以下条款。

A-6.2.6.1 ISO 8211

ISO/IEC 8211 规范是一种用于信息交换的数据描述文件格式。S-100 第 10a 部分规定了设置在记录和字段级别的交换集的结构。它进一步指定了实现它们所需的物理构造的内容,如 ISO/IEC 8211 数据记录、字段和子字段。将记录分组到 ISO/IEC 8211 文件被认为是特定于应用的,因此需要在相关产品规范中进行描述。S-100 仅使用二进制 ISO/IEC 8211 格式。

字段标签:DSID		【Upd】*		字段名称:数据集识别	
子字段名称	标签	格式	子字段内容和说明		
记录名字	RCNM	b11	{10} * *)		
记录标识号	RCID	b14	范围:1 到 2 ³² -2		
编码规范	ENSP	A()	定义编码的编码规范		
编码规范版本	ENED	A()	编码规范版本		
产品标识符	PRSP	A()	产品规范中规定的数据产品的唯一标识符		
产品版本	PRED	A()	产品的版本*		
应用实例	PROF	A()	指定数据产品中实例的标识符		
数据集名	DSNM	A()	数据集名称		
版本号	EDTN	b12	数据集版本号		
更新号	UPDN	b12	数据集更新号		
发行日期	ISDT	A(8)	发行日期格式:根据 ISO 8601 YYYYMMDD		

图 A-6-1 字段示例表

A-6.2.6.2 GML

地理标记语言(Geography Markup Language, GML)是由开放地理空间联盟(Open Geospatial Consortium, OGC)/ISO 19136 定义的一种 XML 语法,用于表示地理要素。GML 既是地理系统的建模语言,也是网络上地理事务的开放交换格式。应当注意的是,GML 中的要素概念是一个非常普遍的概念,不仅包括常规的“矢量”或离散对象,还包括覆盖和传感器数据。具备综合各种形式的地理信息的能力是 GML 应用的关键。

S-100 10b 部分规定了 GML 配置文件作为 S-100 数据产品 GML 应用模式开发的基础标准。GML 应用模式在每一个数据产品中均定义了一个机器到机器的信息交换文件格式,遵循规范的应用模式。

S-100 GML 配置文件定义了用于 S-100 数据产品的 GML 编码的核心 GML 组件,定义了 XML 和 GML 类型的受限子集,排除了 S-100 GML 数据集不需要的 GML 要素。GML 子集用于创建特定的 GML 编码文件。具体方法如图 A-6-2 所示。

A-6.2.6.3 HDF5

HDF5 是 HDF 组开发的一种文件格式,主要用于传输图像和网格数据。HDF5 尤其适用于处理复杂性和扩展性非常重要的数据。任意类型和大小的数据均可以保存在 HDF5 中,包括复杂数据结构和数据类型。HDF5 文件的典型结构如图 A-6-3 所示。

译者注:此处疑原文有误,本译文按译者理解处理。

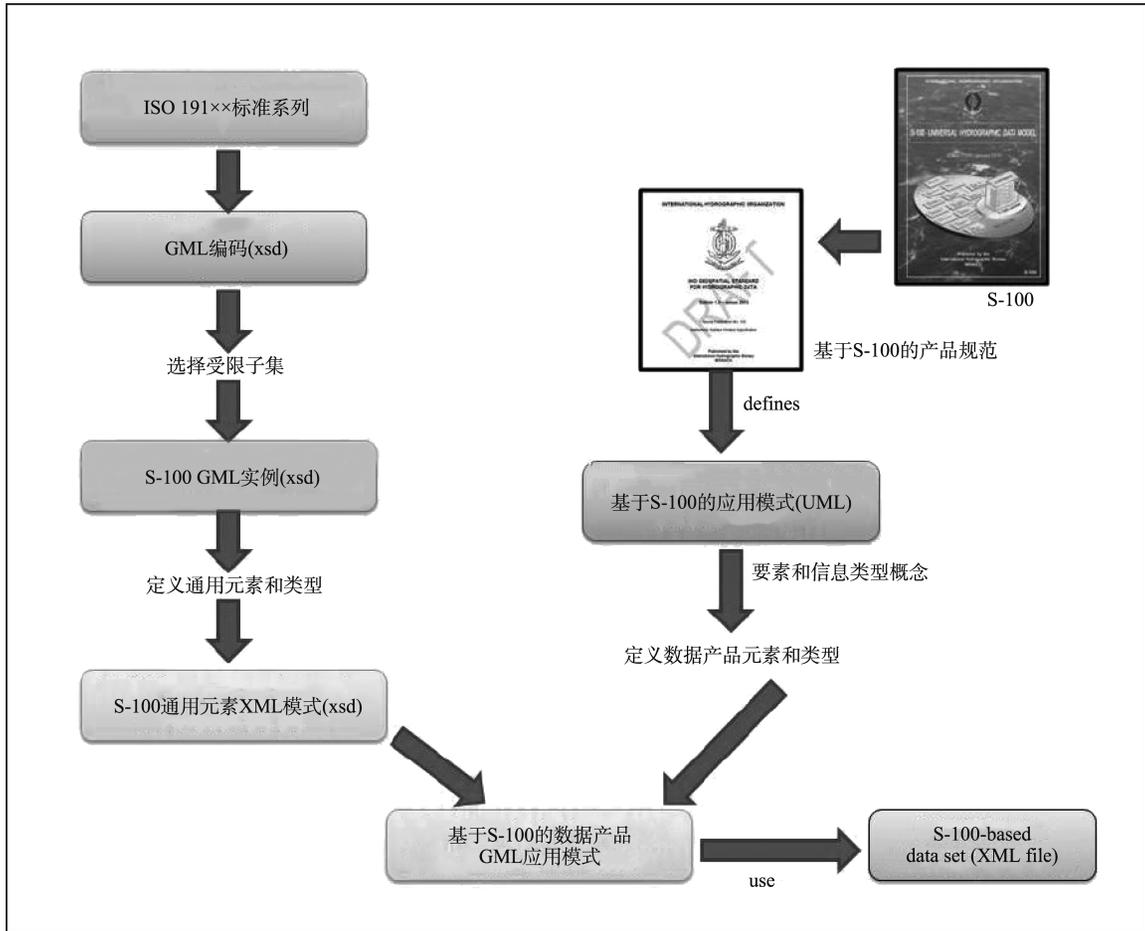


图 A-6-2 GML 专用标准配置文件的继承演化及其在数据产品中的使用

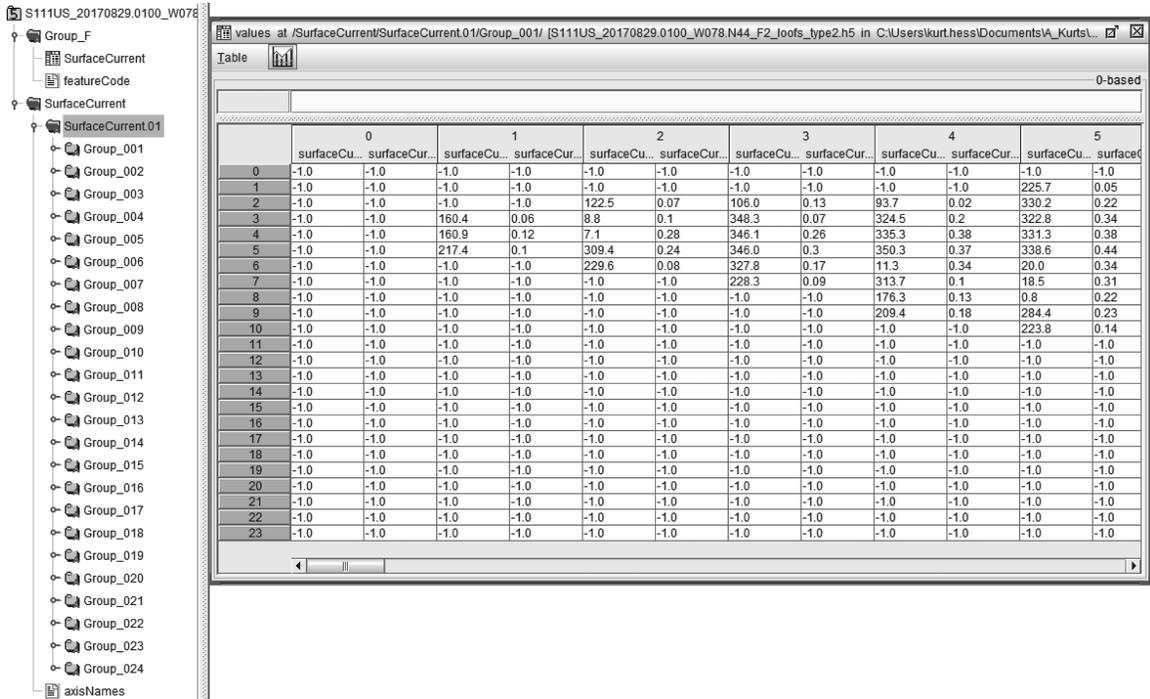


图 A-6-3 显示文件结构 (左侧) 和规则网格数据的二维复合值数组 (右侧)

S-100 第 10c 部分规定了 S-100 采用的 HDF5 专用标准配置文件。它指定了如何以符合 GFM 的方式使用 HDF5,以及如何为 S-100 支持的覆盖类型和点集规定一致的数据格式。

A-6.2.6.4 其他编码和编码标准配置文件

当 S-100 所描述的编码配置文件不满足产品规范的预期用途时,可以在产品规范中指定不同的编码。如果使用的编码不是 S-100 的一部分,应该提供充足的详细信息,以方便制定者理解所选择的编码。

注:如果选择非基于 S-100 的编码,那么支持 S-100 的系统可能无法处理使用该编码的数据产品(例如 ECDIS)。另一种方法是向 S-100 请求添加额外的编码配置文件,并将其添加到符合 S-100 的系统中。

如果使用非 S-100 编码,则必须对元数据进行正确编码,以指示产品规范包含数据产品的编码格式。这是通过每个兼容数据产品的发现元数据中的 S100_DataFormat 属性中的未定义值实现的。详情见 S-100 第 4a 部分。

A-6.2.7 数据内容和结构

本节描述产品规范中使用的不同类型的数据模型元素。S-100 定义了各种不同的元素,但具体允许使用哪些元素取决于各个规范。以下是如何在产品规范中定义元素子集的示例。

示例:

要素类型

<以下条款描述要素目录中可能使用的不同要素类型。>

地理

<地理(Geo)要素类型构成数据集的主要内容,并由其关联的属性和信息类型完全定义。>

元

<元要素包含数据集中其他要素的有关信息。由元要素定义的信息将覆盖由数据集描述性记录定义的默认元数据值。必须最大限度地使用元要素,以便减少单个要素的元属性。>

要素关系

<要素关系将一种要素类型的实例与相同或不同要素类型的实例链接。要素关系有三种常见类型:关联、聚合和组合。>

信息类型

<信息类型是数据集中可识别的信息片段,可以在其他要素之间共享。它们具有属性,但与任何几何图形都没有关系,信息类型可以引用其他信息类型。>

描述产品规范中包含的模型元素的另一个选项是根据某种逻辑方案对元素进行分组,然后描述这些分组。此方法允许类型描述的组合,同时将其与规范中的用法链接。图 A-6-4 是该方法的一个示例,描述了特定产品规范中的所有信息类型。

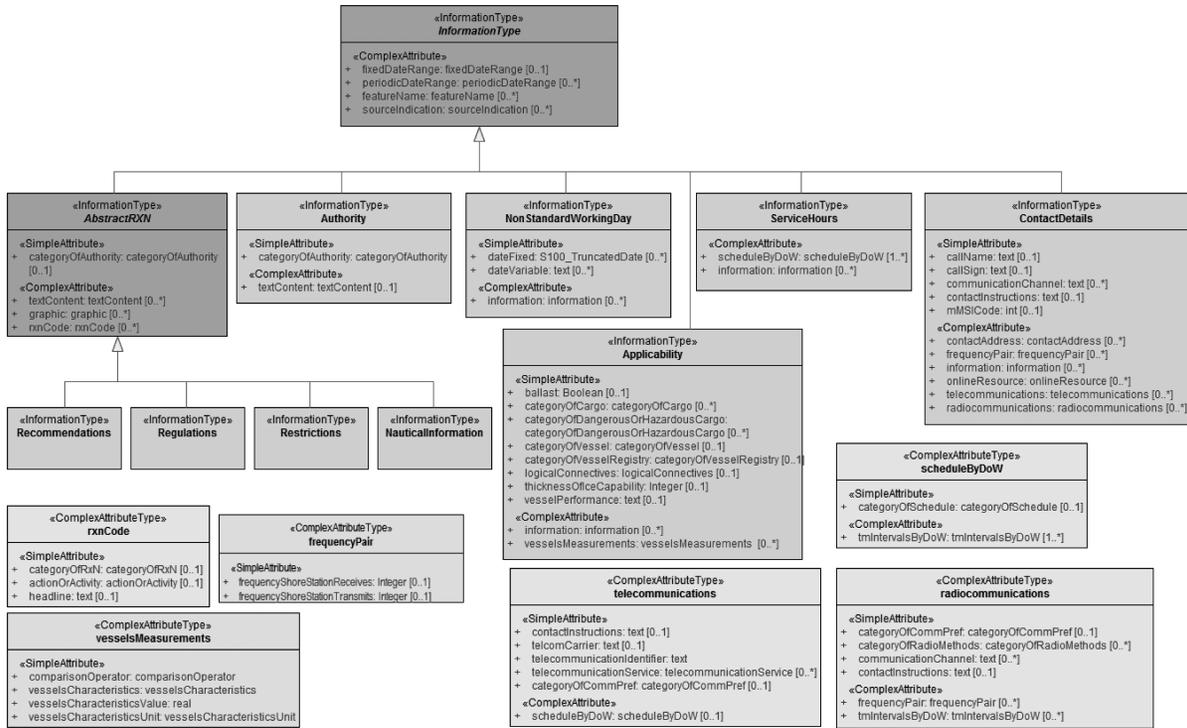


图 A-6-4 信息类型概述示例(S-123)

A-6.2.8 要素目录

ISO 19110 将要素目录定义为包含在一组或多组地理数据中的要素类型、要素属性和要素关联的定义和描述的目录。因此,要素目录充当了机器可读形式的应用模式,并为系统提供了描述符合要素目录的数据集的元素的方法。

当产品规范中的数据模型太大而不能在 UML 中完全重现时,要素目录与整个 UML 应用程序模式的特定子集相结合,可以代替产品规范中描述的完整应用模式的要求。

IHO 在 S-1 × × 编号系统下认可的产品规范必须使用 IHO 的要素目录构建器,以确保符合 S-100 标准。可以在 IHO 要素目录构建器(FCB)之外创建要素目录,但是,要素目录应该根据 S-100 要素目录模式进行验证。

A-6.2.9 数据集

本节指定了数据集的规则。规则的类型可能与数据集大小限制或数据覆盖类型以及数据是否相互重叠有关。

A-6.2.10 数据集加载和卸载

如果数据集在不同的比例尺上有多种表示,则应该考虑加载和卸载方案。应该提供足够的详细信息,以便系统实施人员了解如何以正确的方式创建或加载数据集。例如,任何关于数据集加载的规则都可以使用可视化指令来描述,如图 A-6-5 所示。

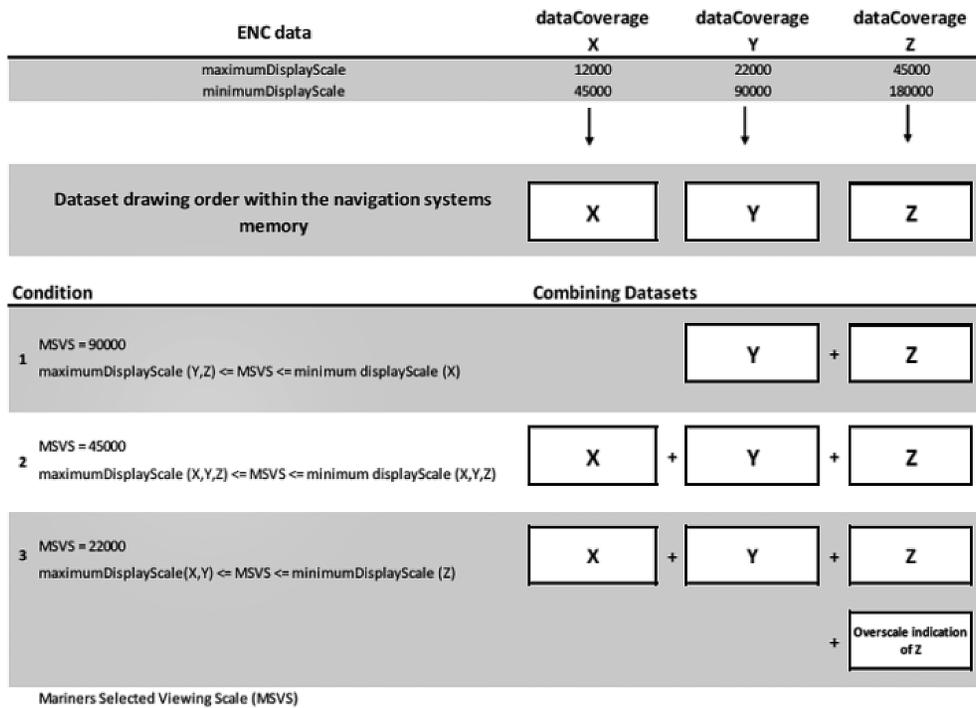


图 A-6-5 S-101 数据加载规则示例

A-6.2.11 几何

本章节描述了几何规则在产品规范中的应用。对于矢量数据,描述中应包括在产品规范中使用 S-100 几何级别的规范,以及 S-100 第 7 部分第 7-4.3 条所述规则的任何例外情况。对于覆盖数据,应该描述使用了 S-100 第 8 部分中的哪个空间模型,例如 TIN 网格数据使用哪种类型的网格,并应包括任何网格、点集或 TIN 结构的规则和特性,见图 A-6-6 中的示例。

A-6.2.12 参考系统

产品规范必须指定数据产品使用的参考系统。所有包含地理参考信息的基于 S-100 的产品规范都应具有水平参考系统,而垂直参考系统适用于具有高度信息或水深数据的产品规范。产品规范可以包括一个以上的垂直参考系统,例如一个用于水深数据,一个用于高程数据。

S-100 的第 6 部分提供了如何描述和指定参考系统的信息。指定参考系统的方法通常是通过在产品规范中建立约定来完成的,该约定通过声明参考系统或使用参考系统列表,然后将信息添加到结果数据集的元数据中。图 A-6-7 是 S-100 产品规范模板的示例。

属性	值	备注
S100-网格覆盖		
维度	2	
起源	网格原点经度 网格原点纬度	来自运营商元数据的值
轴名	“经度,纬度”	
偏移向量	网格间距经度 网格间距纬度	来自运营商元数据的值
程度:底	“0,0”	
程度:高	“numROW-1,numCOL-1”	来自运营商元数据的值
序列规则:类型	“线性的”	
序列规则:启动顺序	“0,0”	
共同点规则	“平均”	
插值类型	“分离的”	对于表面电流没有空间插值
范围类型	名字:数据类型	描述覆盖范围内属性类型的一对; 表面电流速度,单精度浮点数真实
S-100 点覆盖		
域范围		基于所有经度和维度的范围
轴名	“经度,纬度”	
范围类型	名字:数据类型	描述覆盖范围内属性类型的一对
元数据	URI	链接到元数据
共同点规则	名字:数据类型	
几何	GM 点	
值	实数	对应速度和方向值

图 A-6-6 S100_Grid Coverage 和 S100_Point Coverage 的属性及其值示例

水平基准线参考	对从其中获取水平基准值的注册表的引用	1	字符串	EPSG
水平基准值	整个数据集的水平基准	1	整数	4326
垂直基准	整个数据集的垂直基准	1	S100_VerticalAndSoundingDatum	

图 A-6-7 参考系统引用示例

如图 A-6-7 中的示例, EPSG 注册表是一个有用的水平基准注册表。最常用的水平基准 WGS84 的代码列表值为 4326。对于垂直基准和测深基准, S-100 包括一个名为 S100_VerticalAndSoundingDatum 的枚举列表。最常用的垂直基准面和测深基准面包括在这份清单中。可以通过请求 IHO S-100 工作组增加附加值来扩展枚举列表。

A-6.2.13 对象标识符

建议产品规范包含要素和信息对象的持久全局标识符的规则。在现实世界中另有规定的情况下,或已知不需要对该对象引用时,可以省略标识符。例如,无需为制图对象定义标识符。

实例标识符应使用海事资源名称(MRN)的概念和命名空间。MRN命名空间由IALA通过网站<http://mrnregistry.org>进行管理,该网站还包含对适用于MRN概念的全套规则的引用。最顶层的命名空间urn:mrn保持固定,后续的名称空间用冒号分隔,并可通过网站上的申请流程获得。任何希望发布符合MRN标识符的组织都应该从IALA或已经注册了命名空间的组织(如IHO)申请名称空间。S-100第11部分的附件E包含有关MRN概念的其他详细信息。

无论是在产品整个生命周期中持续保留的全局标识符,还是在其他产品中重复使用它们时,均应包含相应的指导信息。在产品之间维护持久的全局标识符有助于实现互操作性,并帮助用户和系统识别数据产品之间的相同要素。

A-6.2.14 数据质量

所有基于S-100的产品规范都应包括全面的获取有关规范所产生的数据质量信息的方法。本指南的C部分对如何解决数据质量方面的问题提供了全面的指导。

A-6.2.15 数据收集和编码说明

任何基于S-100的产品规范都必须提供关于如何收集符合产品规范的数据的信息。这些资料应尽可能详细和具体。为此,S-100产品规范模板建议开发一个数据分类和编码指南(DCEG),用于将真实世界的示例链接到数据模型。例如,DCEG可以解释如何使用特定数据模型对不同类型的水下岩石进行编码,包括应该使用哪个要素类、什么属性以及对应不同类型水下岩石的期望值。

数据收集指南主要由数据生产者使用,它是有关如何根据特定产品规范创建数据的全球通用规则的集体指导文档。随着更多的特殊情况得到解决并进入到全球商定的进程,该文件也将随着经验的增长而完善。这也提高了生产者和产品之间的总体一致性,从而形成更稳定的用户系统,因为所有利益相关方都对如何使用数据产品有了共同的理解。

A-6.2.16 维护

通常,基于S-100的产品规范创建的数据不会无限期地保持有效。因此,需要指定根据产品规范创建的数据的维护方法,包括如何更新数据集和支持文件。

以下是两种主要类型的更新例程:

- 按需更新:这种类型的更新例程意味着在需要时更新数据集,并且在进一步更新之前将其视为当前信息。电子海图和航海出版物是以这种方式维护的两类数据。
- 按计划更新:这种类型的更新例程意味着数据集是按固定的计划或间隔更新的,并且用户总是可以预测新数据集何时可用。表层流和水位信息是以这种方式维护的两类数据。

一旦为根据产品规范创建的产品建立了更新例程,就有必要建立实现这一目标的方法。同样,主要有增量更新和整个数据集替换两种选择,这两种方式将在后面的章节中详细介绍。

A-6.2.16.1 增量更新

增量更新是通过插入新信息或者修正信息来修改数据集的一部分,从而更新先前发布的数据集的方法。如果需要考虑带宽,并且修改的部分在整个数据集的范围内占比相对较小时,此方法非常有用。

例如将两个要素添加到包含数千个其他要素的数据集中,此时,增量更新将是一个比较小的数据集,它只包含主产品数据或基本数据集的修订指令。一旦应用了修订指令,更新后的数据集将包括更新的两个要素。

应该注意的是,对于增量更新,可能会出现这样的情况,即当需要更改的内容过多时,重新发布数据集则更为适用,从而用完全更新的基础数据集替换原始基础数据集,并在此数据集基础上发布新的增量变化。在 S-100 4.0.0 版中,ISO 8211 和 HDF5 编码支持这种类型的更新。GML 编码尚不支持这种类型的编码。

A-6.2.16.2 整个数据集更新

这种类型的更新是用新数据集替换先前发布的数据集。当替换数据改变了先前数据集的全部或绝大部分时,该方法最有意义。例如,当某一自然现象的预测数据被更新的预测数据替换,并且由于数据更新使被替换的数据变为无效时。S-100 中的所有编码都支持这种更新方法。

A-6.2.16.3 支持文件维护

可以通过与支持文件一起使用的元数据来更新基于 S-100 的产品规范中的支持文件。这是通过在发现元数据文件中包含发布日期和管理信息来完成的,该文件包含在更新的支持文件的交换集中。下面是如何为产品规范编写此类指令的示例。

示例:

支持文件的类型在发现元数据的“目的”字段中指示。必须从系统中删除带有“删除”标志的支持文件。当删除或更新指向文本、图片或应用程序文件的要素使其不再引用该文件时,系统软件必须在删除该文件之前检查是否有任何其他要素引用同一文件。

为了避免复杂的管理程序,最好指定每个支持文件在交换集中只使用一次,并将支持文件存储在交换集中的单独文件夹中。

A-6.2.17 数据产品交付

产品规范可以为规范中的每个确定的范围定义交付介质(例如 DVD 或 Web 服务)。这是一个可选的部分,但它包括所交付的数据产品的结构。因此,当数据被交付到包含数据自动化水平的系统时,将标准化的交付结构自动化具有重要意义。指定数据产品何时以不同格式交付(例如 SENC 交付)也很有用。交付到 ECDIS 和类似系统的数据通常需要交换集。S-100 包含了对用于交换地理空间数据及其相关元数据的 S-100 交换集的描述,如图 A-6-8 所示。详情见 S-100 第 4a 部分的附录 D。

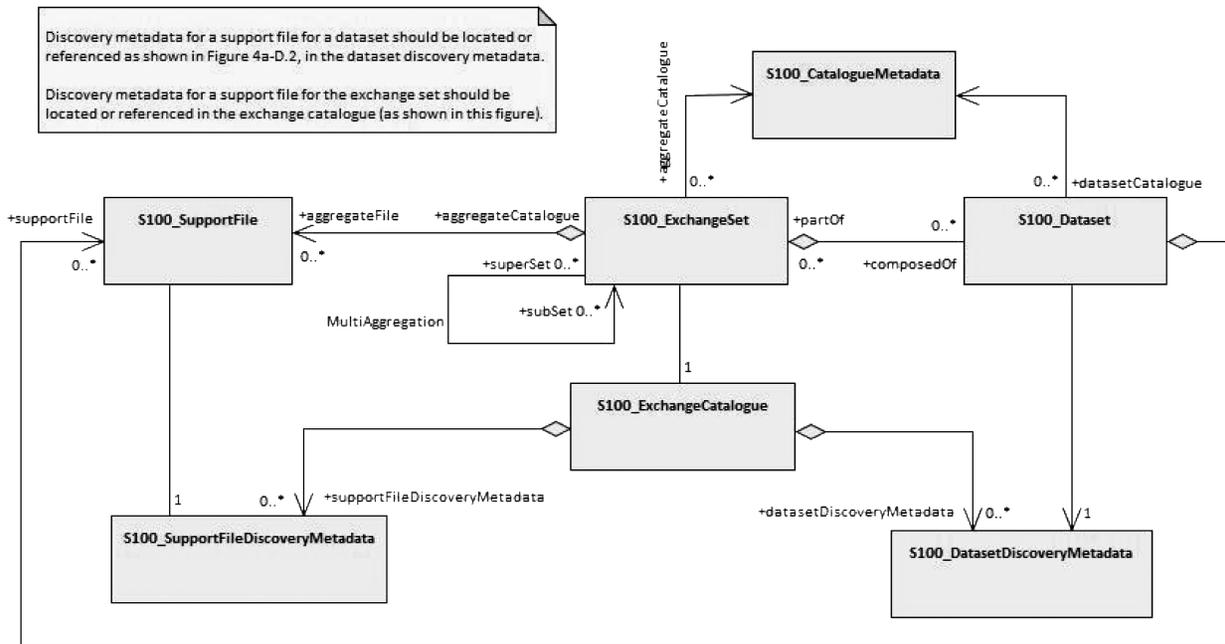


图 A-6-8 S-100 交换集(S-100 4.0.0 版图 4a-D-3)

在交换集中,除了数据集之外,还可以有几种类型的文件。这些类型的文件通常称为支持文件。支持文件可分为两种类型:单个数据集的支持文件和交换集的支持文件。与单个数据集关联的支持文件通常包括的文件类型有文本文件和图像文件,而与交换集相关联的支持文件通常是要素目录和图示表达目录。

根据目标用户的不同,数据产品可以通过多种供应链方式交付,例如通过区域 ENC 协调中心(RENC)、服务提供商、Web 服务、FTP 等。在指定数据产品交付时,应充分考虑供应链方式。

A-6.2.17.1 服务和数据流

S-100 支持一种叫做在线数据交换的替代分发方法,在 S-100 第 14 部分有详细描述。此方法可用于一组数据或具有连续性质的数据,后者也称为“流式数据”,用于动态信息流的交付,例如用在 VTS 中监视船舶轨迹。

应用程序或设备之间的在线数据交换可以遵循不同的通信模式,以支持各种海上业务需求。多个客户端可以与服务交互,以交换 S-100 建模的数据,并可以区分像 AIS 这样的单向消息流和像 Web 要素服务 (WFS) 这样的交互式信息交换。通信的语境可以通过使用面向会话的通信概念实现,即将不同通信伙伴之间的通信分配给一个逻辑实体——会话。这允许为分配给会话的交互定义元数据。

用于服务的通信方式应该在通信栈中定义。指定通信堆栈可确保通信服务的协调一致,且更容易实现。由 ISO-OSI 参考模型定义的堆栈组织的通信:

- 用于定义消息类型的会话协议(例如 WSDL、SOAP、REST、SoS);
- 编码和压缩(例如 GML、XML、ISO8211、HDF 等)完成数据的序列化;
- 通信协议(例如 HTTP)及加密(例如 HTTPS),用于定义网关之间的交互;以及
- 传输层(例如 TCP/IP)及加密(例如 SSL),用于定义网关之间的传输节点。

图 A-6-9 对堆栈进行了描述。

ISO/OSI层

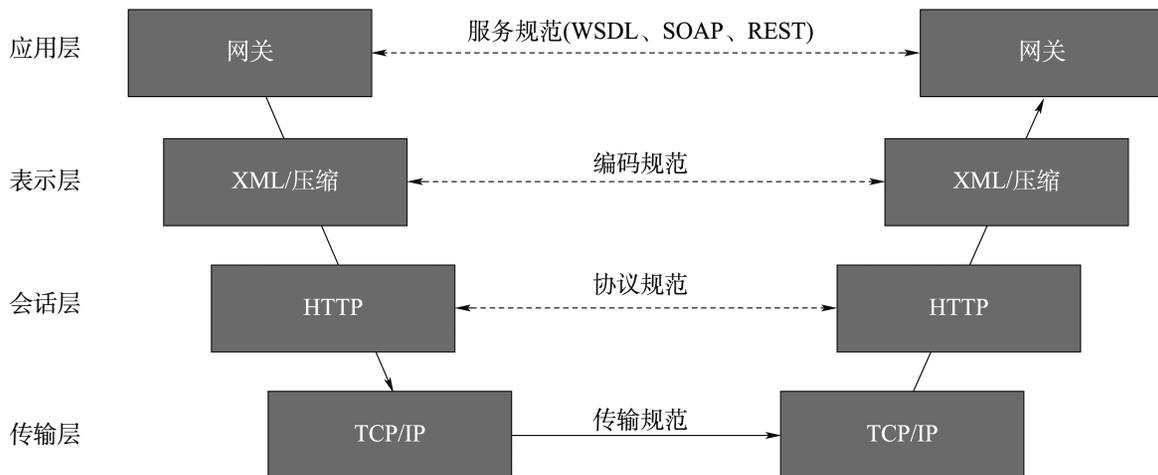


图 A-6-9 通信堆栈

S-100 第 14 部分只涉及应用层和表示层中的概念。下面各层超出了 S-100 的范围。

使用在线交换方法的产品规范必须描述用于构造会话的概念,并且必须解释在会话内传输数据的方式。面向会话的服务通常包含三个组件,每个组件处理不同类型的数据。

- 会话组件:描述会话数据的处理(服务请求、服务响应、登录、登录响应、注销)。

- 服务组件:描述维护服务的信息(例如,保持活动消息、服务状态)。
- 数据组件:描述数据本身,例如,船舶交通图像数据(对象)。

每个组件所需的任何元数据都应在产品规范中详细说明。

A-6.2.18 数据集命名规则

数据集命名应遵循标准模式,以使实现者对即将到来的数据集具有更强的可预见性。建议所有数据集命名约定尽可能遵循这些规则。

XXXYYYY0000000000

XXX 是产品代码(例如 123 是用于海上无线电服务的;101 用于 ENC)。

YYYY 是生产机构代码注册表中的生产机构代码。

0000 是以字母数字字符表示的任意长度的唯一代码。

如有必要,产品规范可以在文件名的“0000”中包含区分字符或代码(例如,下划线(_)字符)。

支持文件应遵循类似的命名规则。

A-6.2.19 元数据

元数据是关于数据的数据。在 S-100 中,元数据的主要目的是提供关于标识的信息、空间和时间范围、质量、应用模式、空间参考系统和分发数字地理数据。它适用于数据集的编目、信息交换活动以及地理和非地理资源的全面描述。

元数据可以满足多种用途,如数据发现、用于联机查看分发和联机参考(URL)、数据使用、数据创建的详细信息、数据拟合度、数据共享、数据管理等。图 A-6-10 描述了元数据的一些典型用途和受众。

S-100 第 4 部分(元数据)规定了一致性数据产品中必须包含的最小元数据元素。此外,S-100 要求发现和质量元数据应分别按照 S-100 第 4a 和 4c 部分进行结构化。特定产品规范所需的任何其他元数据项都必须记录在数据产品规范中,这些元数据项应使用 ISO 19115-1 和 ISO 19115-3 进行定义,并根据需要进行扩展或限制。S-100 第 4 部分的附录 4a-E 包含了扩展或限制最小元数据时适用的规则。

产品规范应用模式应规定如何将元数据打包在符合的数据集中。应在产品规范中为每个确定的范围指定此信息。此外,如果数据产品用于标准化的用户环境(如 ECDIS),则应注意任何显著偏差(例如添加或减少标准元数据)可能导致其在系统中不适用的情况,具体取决于是否已实施了标准的 S-100 模式。因此,在开发产品规范的元数据时要小心,强烈建议保持在 S-100 的元数据范围内。

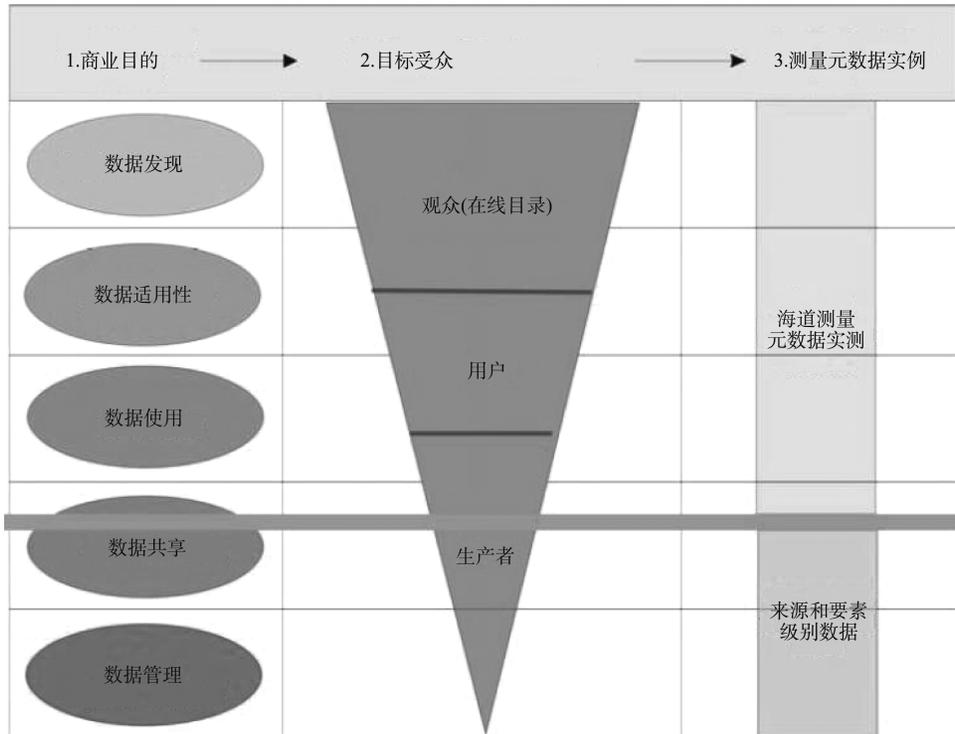


图 A-6-10 元数据层级示例

A-6.2.19.1 发现元数据

对于信息交换,需要几类元数据:关于整个交换目录的元数据、目录中包含的每个数据集的元数据和有关组成包支持文件的元数据。这些数据在 S-100 中称为发现元数据,在交换集中使用发现元数据,可以使用户在不必打开每个数据集或支持文件的情况下就能够了解相关内容。

A-6.2.19.1.1 数据集的发现元数据

S-100 指定数据集的发现元数据包含在根据 S-100 元数据模式创建的外部 XML 文件中。此元数据集符合最小元数据,并在一些地方进行了扩展,以提供更多详细信息(例如,有关参考数据和数据发布日期的信息)。有关数据集的发现元数据的更多信息,请参见 S-100 规范附录 4a-D 信息交换目录的发现元数据。

A-6.2.19.1.2 支持文件的发现元数据

S-100 指定支持文件的发现元数据包含在根据 S-100 元数据模式创建的外部 XML 文件中。此元数据集符合最小元数据,并对其进行扩展,以提供有关支持文件管理的信息,以便对其进行更新。更多详情见 S-100 规范附录 4a-D 信息交换目录的发现元数据。

A-6.2.19.1.3 数据流和服务的元数据

S-100 规范附录 4a-D 信息交换目录发现元数据中也描述了数据流或服务的元数据。S-100 第 14 部分规定了附加元数据和其他元数据。到 4.0.0 版为止,这两个版本还没有一致。此时,产品规范开发人员应使用第 14 部分的元数据,并根据需要补充第 4 部分的元数据,包括使用第 4 部分附录 4a-E 中详述的元数据扩展原则。

A-6.2.20 图示表达

图示表达是产品规范的可选部分。但是,如果跨所有用户平台的一致图示对于基于 S-100 的数据产品很重要,那么指定如何完成图示表达将变为强制性要求。在 S-100 产品规范中,这部分是通过包括一个图示表达目录来完成的。图示表达目录是要素目录的已定义图示指令的集合,包括图示功能、符号和图示语境。

S-100 中可能有 XSLT 和 Lua 两种类型的图示表达目录。S-100 的第 9 部分提供了有关产品规范如何包含从提供的抽象模式派生的输入模式的说明,一组映射规则(在 XSLT 或 Lua 中定义)、一组符号(以 SVG 格式定义)、线样式、颜色等使其可用于产品数据集。可以通过多种方式创建图示表达目录,包括手动和使用图示表达目录构建器(了解更多详细信息,参见本指南第 7.4 条)。

例如,图示表达目录可以在交换集中提供,并且可以与要素目录和数据集组合。确切的分发方法应在产品规范中定义,但应考虑有效分配,并尽可能减少数据量。因此,考虑某种形式的图示表达目录的集中分发可能是有益的。

产品规范应包括使用图示表达目录的实施者的说明,包括使用数据的语境。

许多 IHO 产品规范将用于具有一定型式认证要求的系统(例如 ECDIS)。船级社进行型式认证的说明应酌情添加到产品规范中。这些说明应包括微小偏差容许度指南和微小偏差的定义。

A-6.2.21 附加信息

产品规范应包含所有的详细信息,以便让预期的利益相关者更容易实施。但是,可能还需要考虑到实施者、用户和其他利益相关者的其他因素。这些额外的注意事项可以添加到实施指南的章节或附录中,或者添加到另一个适当的标题中。这一节可用于说明预期用途的背景,或详细说明可能影响利益相关者的特殊情况等。

A-7 IHO S-100 基础设施

本节介绍为支持 S-100 框架和 e-航海概念而开发的 IHO 基础设施。

A-7.1 IHO 地理空间基础设施(GI)注册系统

注册系统是注册表集合所在的整个信息系统(或位置)。注册表是数据库中表的集合,其中包含分配给具有相关项说明的项的标识符。描述可能包含许多类型的信息,包括名称、定义和代码。

在 S-100 中,IHO 托管了一个名为 S-100 地理空间信息(GI)注册表的在线注册系统引擎。可在此处访问:<http://registry.iho.int>。此注册系统提供访问和维护各种 S-100 注册表的工具。S-100 GI 注册表包含次级注册表,如图 A-7-1 所示。

每种注册表类型可进一步细分为若干领域,例如海道测量、内陆 ENC、助航信息等。注册表由 IHO S-99 S-100 地理空间信息注册系统的组织和管理业务程序负责管辖。

注册系统和注册表概念的主要优点是其灵活性,允许在概念注册表中使用独特的标识和分类对类似条目进行多个版本的注册。条目分为以下几类:

- 有效(最新版本);
- 取代(以前的版本);
- 取消(不再建议使用);或
- 无效(提议但未被接受或不再被接受)。

由于此种分类和时间戳,要素目录的某一个版本引用的项始终是合法的,即使被引用项在稍后日期注册了较新版本。这意味着,如果注册了新产品或修改了现有产品,则不会因此自动要求提供新版本的产品规范。注册系统中专门列入了“无效”项目的类别,以帮助标识不适用的项。



图 A-7-1 注册系统(第3版)

A-7.2 要素目录构建器

要素目录是产品规范数据模型的机器可读表达式。它既可以用现成的 XML 编辑器构建,也可以用要素目录构建器(FCB)构建。这两种情况都必须符合 S-100 第 5 部分的结构和 S-100 要素目录模式。IHO 提供了一个 FCB,任何人都可以使用它为基于 S-100 的产品规范创建一个要素目录。该软件与 IHO GI 注册表交互,提供一种机制,用于将注册表中可用的元素绑定在一起,以形成要素、属性、枚举列表及其可用值等等,图 A-7-2 说明了 IHO 要素目录构建器的概念。

FCB 还包含一个用于处理尚未添加到注册表中的提案的功能。

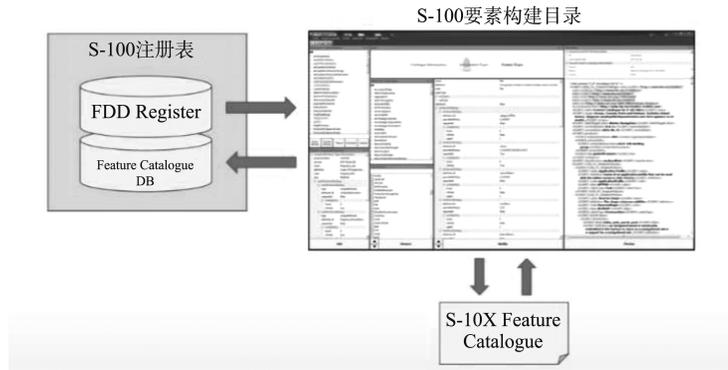


图 A-7-2 S-100 要素目录构建器

A-7.3 DCEG 构建器

为了简化产品规范的数据分类和编码指南(DCEG)的创建,IHO 开发了 DCEG 生成器。该工具利用要素目录为 DCEG 表创建绑定和输入。然后,将图像和特定文本手动添加到 DCEG 表的编码部分。图 A-7-3 显示了通过 DCEG 构建器创建 DCEG 的过程的概述。需要应用模式来编制要素目录,这等同于应用模式的 XML 表示形式。要素目录的使用可以在 DCEG 中描述。

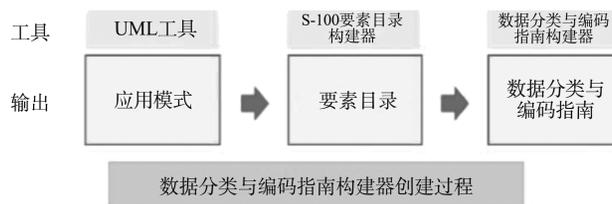


图 A-7-3 使用 DCEG 创建 DCEG 的过程

在准备产品规范时,DCEG 构建器的使用是可选的。任何基于 S-100 的产品规范都必须包含如何以符合规范的方式对信息进行编码的说明,但这些说明不必遵循特定的 DCEG 样式。然而,DCEG 样式很容易理解,并且通过使用 DCEG 构建器,还很容易创建要素和信息类型、属性、关联和编码实例的表格²,并确保这些表格与要素目录一致。IHO 样式的要素表(或等效要素表)可以使用普通的 Office 文字处理软件来准备和维护,但经验表明,确保最初和持续符合 XML 要素目录可能是一项重要任务,需要花费大量精力来维护和保持最新。图 A-7-4 显示了 DCEG 构建器是如何连接到 GI 注册表的,并显示 DCEG 生成器如何在创建要素目录的过程中分级。使用构建器可以极大地简化开发过程并提高与要素目录的一致性。应当注意,DCEG 构建器依赖于在要素目录数据库中注册的要素目录。

² 此类表格的建议格式见 IHO S-100 产品规范模板,该模板在 S-100 第 11 部分附录 11-D 中引用,可从 IHO 网站获得。

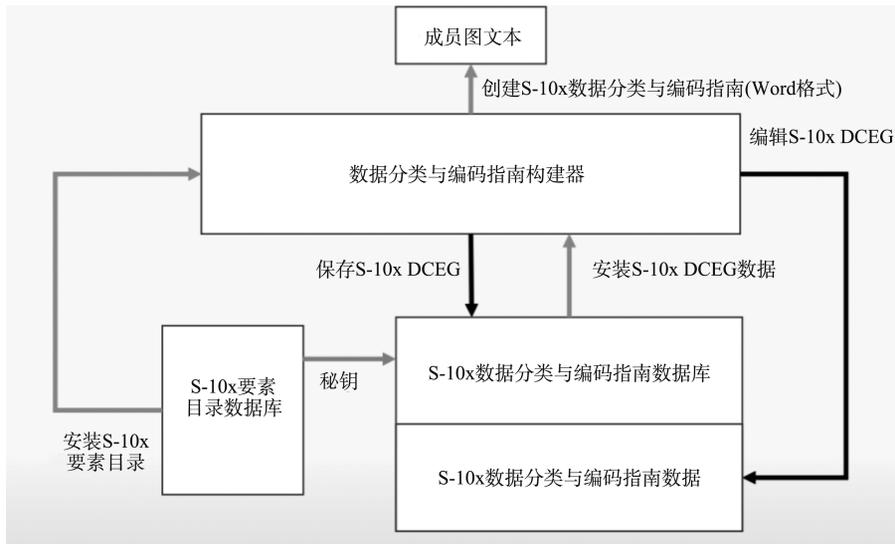


图 A-7-4 DCEG 构建器架构

A-7.4 图示表达目录构建器

图示表达目录是机器可读的指令,用于说明如何描述符合特定产品规范特定版本的数据模型。它们可以手动构建,也可以通过图示表达目录构建器(PCB)创建。在任何情况下,它们必须符合 S-100 第 9 部分规定的结构和 S-100 图示表达目录模式。IHO 基础设施包含一个用于 XSLT 的 PCB,希望产品规范开发团队利用它为基于 S-100 的产品规范创建图示表达目录。

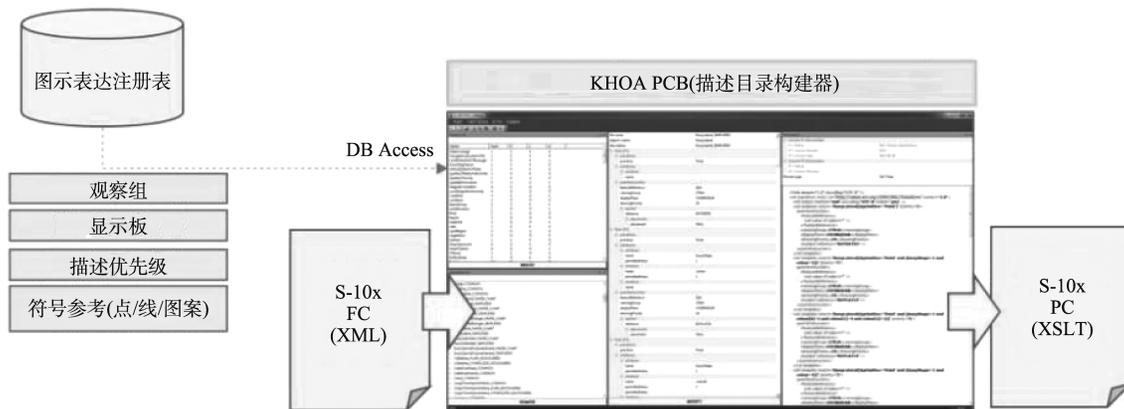


图 A-7-5 图示表达目录构建器

该软件与 IHO GI 注册表和要素目录中的图示表达注册表交互,并提供一个接口,用于将注册表中可用的元素绑定在一起,以形成要素目录中所需元素的符号、线样式和区域形式,图 A-7-5 说明了图示表达目录构建器的概念。

LUA 图示表达目录的 PCB 计划在未来提供。

B 部分 编制产品规范的步骤

B-1 概述

S-100 通用海道测量数据模型是一种海道测量地理空间数据标准,可以支持各种海道测量相关的数字数据源,并与主流国际地理空间标准,特别是 ISO 19000 系列地理标准完全一致。这种方式使海道测量数据和应用程序更容易集成到地理空间解决方案中。S-100 本质上比 S-57 更灵活,并制定了使用图像和网格数据类型、增强的元数据和多种编码格式的条款。它还通过专门的在线注册表为要素、属性和图示提供了更灵活和动态的维护机制。S-100 提供了一个组件框架,其能够为海道测量数据建模建立标准化的产品规范,从而在不同的数据标准和系统之间提供真正的互操作性。

S-97 是一本指南,目标读者是基于 IHO S-100 标准(通用海道测量数据模型)产品规范的开发人员和维护人员。

B-2 介绍

B 部分旨在为开发基于 S-100 的产品规范者解释产品规范的各个组件。它描述了典型的产品规范的典型开发阶段,以及每个阶段的规定顺序。它为如何使用 S-100 中包含的规则来构建产品规范组件提供了指导,包括文档组件、UML 模型以及机器可读组件(如 XML 编码的要素目录和数据格式)。

B-3 规范性引用文件

ISO 8211 ISO 8211 信息交换结构实现的数据描述型文件规范。ISO/IEC 8211,1994 年。

ISO 19109 地理信息 应用模式规则,ISO 19109:2005。(由 ISO 19109:2015 取代,但 S-100 4.0.0 版仍参考 2005 版。)

ISO 19115-1 地理信息 元数据 第 1 部分:基础,ISO 19115-1,2014 年,2018 年第一次修正。

ISO 19115-2 地理信息 元数据 第 2 部分:图像和网格数据扩展,ISO 19115-2,2009 年。

ISO 19115-3 地理信息 元数据 第 3 部分:基础概念的 XML 模式实现,ISO/TS 19115-3,2016 年。

- ISO 19119 地理信息 服务,ISO 19119,2016 年。
- ISO 19136 地理信息 地理标记语言(GML),ISO 19136,2007 年。[也可作为 OGC 07-036 地理标记语言(GML) 编码标准提供。开放地理空间联盟公司,2007 年。]
- ISO 19157 地理信息 数据质量,ISO 19157,2013 年,2018 年第一次修正。
- ISO 19757-3 信息技术 文档模式定义语言(DSDL) 第 3 部分:基于规则的有效性验证 图创。IHO 决议第 2/2007 号(修订)修改国际海道测量组织技术标准和规范的原则和程序。
- S-52 ECDIS 海图内容与显示规范。
- S-58 IHO S-58 电子海图有效性检查,6.1.0 版,2018 年 9 月。
- S-99 IHO S-99 S-100 地理空间信息注册系统组织与管理操作程序,1.1.0 版,2012 年 11 月。
- S-99A IHO S-99 附录 A IHO GI 注册系统内容惯例和指南。(正在开发中)
- S-100 IHO S-100 通用海道测量数据模型,4.0.0 版,2018 年 12 月。
- S-101 IHO S-101 电子海图产品规范,1.0.0 版(草稿),2018 年 7 月。
- S-111 IHO S-111 表层流产品规范,1.0.0 版(草稿),2018 年 6 月。
- S-122 IHO S-122 海洋保护区,1.0.0 版,2019 年 1 月。
- S-123 IHO S-123 海上无线电服务,1.0.0 版,2019 年 1 月。
- SVGTiny 可扩展矢量图形(SVG)简洁版 1.2 规范,W3C 推荐标准,2008 年 12 月。
URL: <<http://www.w3.org/TR/2008/REC-SVGTiny12-20081222/>>
- 注:除非特别标明版本,本文档提及的 S-100 均指 S-100 4.0.0 版。

B-4 术语和缩略词

B-4.1 术语

抽象类 **abstract class**

无法实例化的对象类,或在信息模型中被指定为不允许实例化的对象类。

注:抽象类的子类可以是抽象的,也可以是非抽象的。

聚合 **aggregation**

一种特殊的关联形式,用于指定聚合体(整体)和组成部分(见组合)之间的整体与部分关系。[ISO 19103]

应用 application

支持用户要求的数据操作和处理。[ISO 19101-1:2014]

应用模式 application schema

一个或多个应用所需数据的概念架构。[ISO 19101-1:2014]

关联 association

两个或多个类之间的语义关系,用于指定它们的实例之间的连接关系。[ISO 19103]

属性 attribute

(1)实体的命名特性。[ISO/IEC 2382-17:1999]

注:描述实体的几何、拓扑、主题或其他特征。

(2)UML:分类器中要素实例可能包含的值范围。[ISO/IEC 19501:2005(改编)]

(3)XML:元素中包含的名称与值对。[ISO 19136]

base64

一种用于表示任意序列的二进制八位数编码,其格式允许同时使用大小写字母,但不必是人类可读的。

[IETF RFC 4648(重新设计)]

代码 code

遵循特定方案的表示标签。[ISO 19118:2011]

代码表 codelist

包含允许值代码的值域。[ISO 19136]

代码空间 codespace

代码、名称、术语或类别的规则或权限。[ISO 19136]

示例:代码空间的示例包括字典、权威机构、代码表等。

组合 composition

作为整体一部分的具有强所有权和相同生命周期的聚合关联形式。[ISO 19103]

概念模型 conceptual model

定义话语世界概念的模型。[19101-1:2014]

概念模式 conceptual schema

概念模式的正式描述。[ISO 19101-1:2014]

数据客户端 data client

接收基于 S-100 的加密数据的终端用户。数据客户端将使用软件应用程序(例如 ECDIS)来执行 S-100 保护方案中详细描述的诸多操作。[S-100 第 15 部分(改编)]

例如:ECDIS 用户。

数据许可 data permit

包含解密许可产品所需的加密产品密钥的文件,通常是专门为特定数据客户端创建的。[S-100 第 15 部分(改编)]

数据服务器 data server

生产加密数据文件或向数据客户端颁发数据许可证的组织。[S-100 第 15 部分(改编)]

要素 feature

真实世界现象的抽象表现。[ISO19101:2003]

注:要素可能以类型、类或实例的形式出现。要素类型或要素实例仅用于一种情况时应予以明确。要素类应在模型或应用程序架构的上下文中使用。

示例:名为“埃菲尔铁塔”的真实世界现象可以与其他现象一起归入名为“铁塔”的要素类型。

要素关联 feature association

将一种要素类型的实例与相同或不同要素类型的实例链接的关系。[ISO 19110]

要素属性 feature attribute

要素的特征。[ISO 19101]

注:一个要素属性类型有一个名称、一个数据类型和一个与其相关联的域。要素属性实例具有从要素属性类型的值域获取的属性值。

示例 1:名为“颜色”的要素属性可能具有属于数据类型“文本”的属性值“绿色”。

示例 2:名为“长度”的要素属性可能具有属于数据类型“实数”的属性值“82.4”。

要素目录 feature type

包含定义和描述一组或多组地理数据中出现的要素类型、要素属性和要素关联的目录。[ISO 19110]

要素类 identifier

应用模式或模型中表示要素的类。

要素类型 information type

要素目录中对要素、属性及关联进行描述的元素。

标识符 identifier

一种语言上独立的字符序列,能够唯一和永久地识别与之相关的字符。[改编自 ISO/IEC 11179-3:2003]

信息类型 information type

仅具有主题属性的数据集中可识别的信息单元。[S-100 3-5.1.2(改编)]

实例化 instantiate

用具体实例表示。[韦氏词典线上版 <https://www.merriam-webster.com/dictionary/instantiate>]

接口 interface

描述实体行为的命名操作集。[ISO 19119:2005]

元数据 metadata

有关资源的信息[ISO 19115-1];定义和描述其他数据的数据。[ISO 11179-3:2013]

模型 model

现实世界某些方面的抽象。[ISO 19109-2015]

操作 operation

可以调用对象来执行的转换或查询的规范。[ISO 19119:2005]

注:一个操作有一个名称和一个参数列表。

注册表 register

一组文件,其中包含分配给具有相关项描述的项的标识符。[ISO 19135]

注:描述可能包括许多类型的信息,包括名称、定义和代码。

注册系统 registry

维护注册表的信息系统。[ISO 19135]

关系 relationship

模型元素之间的语义联系。[ISO 19103]

资源 resource

符合要求的可识别资产或手段。[ISO 19115-1]

示例:数据集、数据集系列、服务、文档、计划、软件、人员或组织。

方案管理员 scheme administrator

专门负责维护和协调 S-100 规定的保护方案的组织。[S-100 第 15 部分(改编)]

服务 service

实体通过接口提供的功能的不同部分。[ISO 19119:2005]

空间对象 spatial object

用于表示要素的空间特性的对象。[ISO 19107:2003]

流 stream

在线数据交换中,由通信系统传送的连续的碎片化数据序列。[S-100]

话语范围 universe of discourse

对包括所有感兴趣的事物的现实或假想世界的看法。[19101-1:2014]

词汇 vocabulary

包含一个或多个特定主题领域的名称和定义的术语词典。[ISO 1087-1:2000]

B-4.2 缩略词

- CORBA 通用对象请求代理体系结构(Common Object Request Broker Architecture)
- COTS 商用成品(Commercial Off-The-Shelf)
- CRS 坐标参考系统(Coordinate Reference System)
- CSS 层叠样式表(Cascading StyleSheets)
- DCEG 数据分类与编码指南(Data Classification and Encoding Guide)
- ECDIS 电子海图显示与信息系统(Electronic Chart Display and Information System)
- ECS 电子海图系统(Electronic Chart System)
- ENC 电子海图(Electronic Navigational Chart)
- FC 要素目录(Feature Catalogue)
- GFM 通用要素模型(General Feature Model)
- GI 地理空间信息(Geospatial Information)
- GML 地理标记语言(Geography Markup Language)
- HSSC 国际海道测量组织服务和标准委员会(IHO Hydrographic Services and Standards Committee)
- IALA 国际航标协会(International Association of Lighthouse Authorities)
- IEC 国际电工协会(International Electrotechnical Commission)
- IETF 因特网工程任务组(Internet Engineering Task Force)
- IHO 国际海道测量组织(International Hydrographic Organization)
- IMO 国际海事组织(International Maritime Organization)
- ISO 国际标准化组织(International Organization for Standardization)
- JCOMM 海洋学和海洋气象学联合技术委员会(WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology)
- MRN 海上资源名称(Maritime Resource Name)
- OEM 原始设备制造商(Original Equipment Manufacturer)
- RENC 区域电子海图协调中心(Regional ENC Coordinating Centre)
- SVG 可扩展矢量图形(Scalable Vector Graphics)
- W3C 万维网联盟(World Wide Web Consortium)
- XML 可扩展标记语言(eXtensible Markup Language)
- XSD XML 模式定义(XML Schema Definition)
- XSL 可扩展样式表语言(eXtensible Stylesheet Language)
- XSLT XSL 转换(XSL Transforms)

B-5 开发步骤概述

在开发基于 S-100 的产品规范时,建议开发人员考虑基本开发过程中定义的每个阶段。本节总结了产品规范编制的各个阶段。接下来的部分将更详细地描述这些阶段。

B-5.1 基本开发过程

下面的描述适用于新的数据产品规范开发。

- 1) **启动**: 确定对新数据产品的需求, 定义其范围, 并确定新产品与现有数据产品规范之间的界限, 获取样品源材料, 描述典型的应用场景。
- 2) **开发域模型/应用模式**: 定义那些可以描述域和与数据产品相关的类和属性。定义类之间的关系并指定适用的约束。准备一个或多个描述域模型的 UML 图。
- 3) **从域模型填充 IHO GI 注册系统**: 提出对 IHO GI 注册系统中建议修改的现有的类和属性, 并建议使用系统接口将新的类和属性添加到 GI 注册表中。跟踪从注册系统管理员或域控制机构返回的任何建议或询问。
- 4) **开发要素目录**: 使用要素目录构建器, 根据 GI 注册表中批准的要素和信息类、属性和关系准备 XML 要素目录。
- 5) **定义打包和数据传输模式**: 确定是否通过 Web 服务(如果是, 标识或概述服务协议)、电子邮件等将数据产品作为传输(交换)集交付数据文件。确定数据是实时传递还是近实时传递。确定由交付机制和通信约束(如消息大小、带宽限制、客户通信可用性、许可和授权、支付和加密等)引起的约束和要求。
- 6) **定义元数据**: 调查 S-100 中列出的元数据元素是否适合数据产品及其可用的打包和交付方法。为 S-100 中列出的元数据元素定义适当的值和限制。考虑是否需要其他特定产品的元数据元素。
- 7) **定义数据格式**: 选择适当的数据格式。S-100 提供了 3 种标准交付格式(ISO 8211、GML 和 HDF5)。定义强制格式的项(例如, 嵌入的标头元数据)。如果需要, 准备特定格式的制件(例如, GML 格式的 GML“应用模式”XSD 文件)。
- 8) **开发数据分类和编码指南(DCEG)**: DCEG 面向制图员、编辑和数据编码人员, 而不是应用程序开发人员或原始设备制造商, 应该从这个角度编写。DCEG 应包含足够的关于基本概念(如数据类型、要素、信息类型和关联等)的概览和一般材料, 以使其目标受众对需要应用的概念有一个基本的认识, 它的主要内容应该是编码内容以及如何对其进行编码。

- 9) **定义图示表达符号和规则**:确定用于图示表达的符号以及从数据产品生成显示的规则。这意味着决定应该显示的要素、属性值或关联数据对象的组合应分配不同的符号,以及预先存在的符号和颜色哪些可以重用,哪些需要定义。它还包括以适当的格式定义新元素。
- 10) **图示表达元素的注册**:使用 IHO GI 注册系统界面向 GI 注册表中的图示表达注册表中提供新的图示表达组件(例如符号、颜色标志、线样式和区域填充等)提案。跟踪从注册系统管理员或域控制机构返回的任何建议或查询。
- 11) **开发图示表达目录**:为预期应用程序域和使用方案中显示的要素和信息类型准备一个或多个图示表达目录,包括在步骤 8)中确定的符号,并以适当的图示表达语言编码相应的规则。
- 12) **定义空间参考系统**:确定推荐的坐标参考系统和垂直基准。
- 13) **定义数据产品打包和维护方法**:定义交付包的内容和结构,数据的更新以及随数据交付或作为数据附件交付的任何辅助内容。
- 14) **定义有效性检查和质量度量**:定义对数据集的空间、结构和概念完整性的测试方案。定义有效性检查的特定格式实现方法[例如,GML 格式的 Schematron 规则(ISO 19757-3)]。为数据集定义质量指标(质量指标通常是关于“通过测试”和“失败测试”获得的统计数据)。
- 15) **确定与其他数据产品的互操作性**:与 IHO 互操作性目录维护小组一起,确定互操作性目录中的任何产品组是否因新数据产品得到补充或增强,并确定 IHO 互操作性目录将如何受到新产品的影响,包括显示优先级、交错、预定义组合以及其他互操作性规则和操作的更新。
- 16) **为测试平台准备样品数据**:创建符合产品规范中定义的数据格式、打包和要素目录的示例数据集和交换集。
- 17) **测试和反馈**:测试数据生产和样本数据在特定应用系统中的使用,以验证产品规范的正确性、完整性、一致性和实用性,包括相关制件,如要素目录和 XML 模式。

产品规范的开发是一个不断完善的迭代过程,按照 S-100 成熟度阶段进行推进。开发者应该根据每个阶段的经验和结果,以更小或更大的周期多次循环执行以上阶段。例如,样本数据的第一次准备(阶段 16)可能识别域模型中的差距,并且需要重新访问阶段 2 的域模型的开发,或者图示表达规则的开发(阶段 9 ~ 11)可以识别域模型中的差距(阶段 2)。在第二次及以后的过程中,迭代过程的部分阶段不再重要,因为结果可能不受后期阶段发现的影响。例如,打包和交付或空间参考系统的选择可能不受样本数据编制过程中发现的变化影响。

图 B-5-1 描述了整个过程,也包括预期的迭代过程。

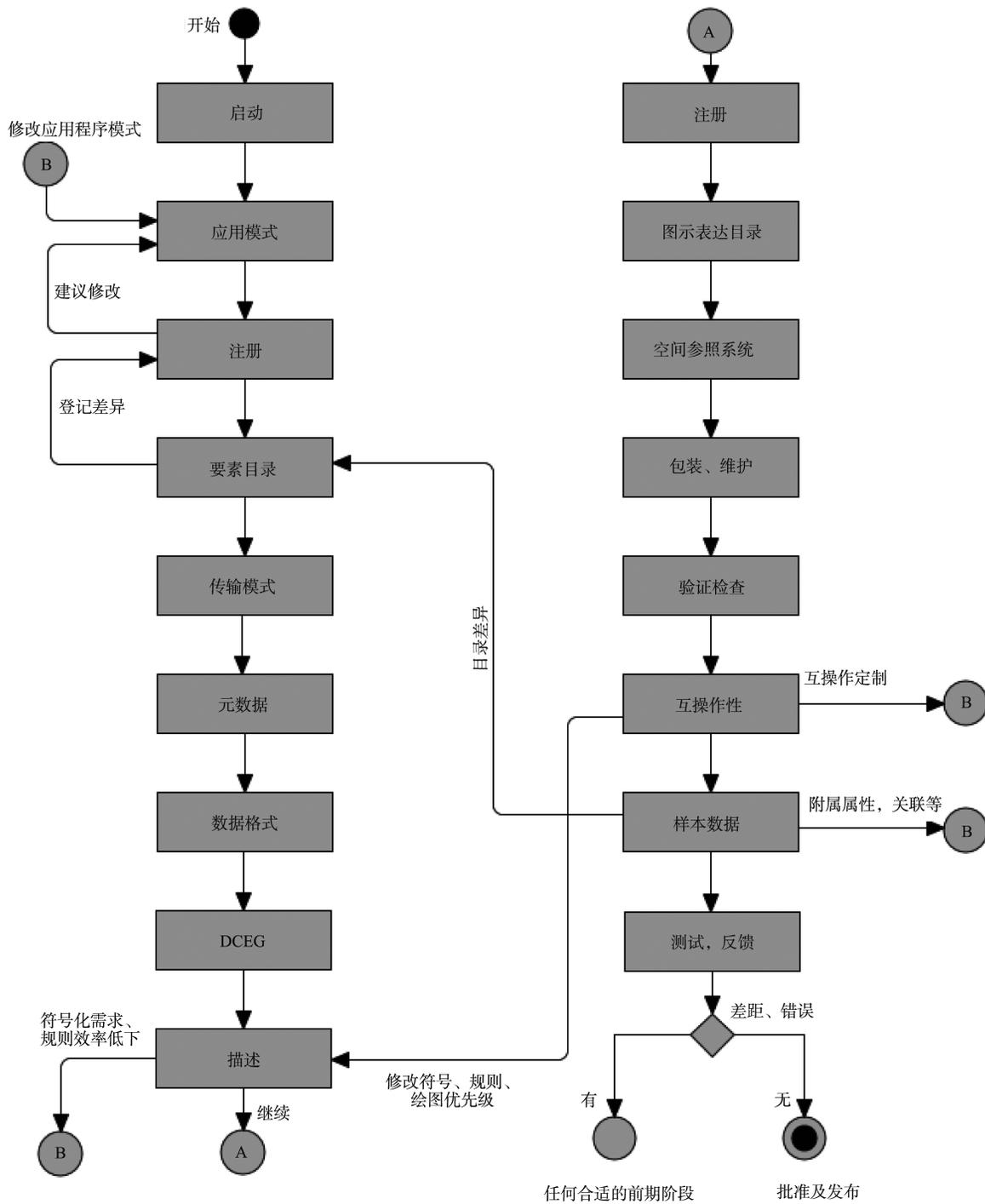


图 B-5-1 产品规范编制过程

S-100 第 11 部分附录 11-A 描述了理想的开发域模型和相关注册表操作的基本步骤过程(图 11-A-1 产品规范过程)。在过程的最后,该图还包括“坐标参考系统”和“产品规范文档”的黑盒阶段。实际过程可能更多地是一个迭代求精的过程。

B-5.2 基于 S-100 的产品规范的审查周期

IHO 第 2/2007 号决议(经修订)概述了基于 IHO 的产品规范的开发和批准过程。作为生命周期的一部分,工作组和项目团队应与 IHO 数据质量工作组(DQWG)合作,审查产品规范是否符合 S-100 对数据质量方面的要求和完整性。虽然此审查可以在 1.0.0 版(测试和实施版)之前进行,但 DQWG 将在 1.0.0 至 1.9.n 版本期间进行审查并提供反馈。任何基于 IHO S-100 1.0.0 版的产品规范都不是常规使用版本,不以提供运营服务为目的。产品规范的 2.0.0 版被认为是成熟的,可以用于实施。图 B-5-2 显示了 IHO 国际海道测量组织第 2/2007 号决议所记录的 IHO 组织审查周期。

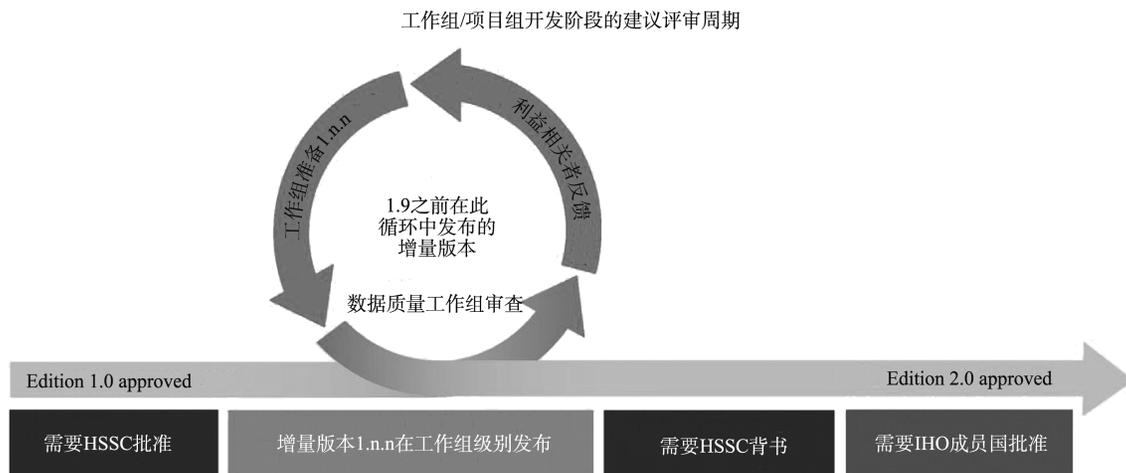


图 B-5-2 审查周期[IHO 第 2/2007 号决议(经修订)]

B-5.3 规范扩展过程

如果证明需要扩展现有的产品规范,开发者应审查第 5.1 节中基本过程的步骤,并确定哪些步骤可以省略。

例如,如果要添加、删除或修改要素、信息类型、属性或关系,则需要基本流程中的大部分步骤。在这种修订中,空间参考系统和打包/维护可能不会改变,因此相应的阶段可以跳过。

如果扩展包括定义新的传输模式,例如为最初打算在交换集中传输的数据产品添加流式传输模式,则可能只需要执行以下阶段:

- 传输模式:要添加传输的流模式,应定义服务器的规范等。
- 元数据:定义交换集模式中哪些元数据元素应用于流模式,是否需要任何新的元数据,以及如何在新模式中交付或提供元数据。

- 数据格式:考虑已经定义的数据格式是否适合于流式传输,是否与预期的服务交付和协议兼容,并在必要时指定适当的格式。
- 打包和维护:指定适当的数据包装格式和服务协议,例如 WFS、REST 等。允许定义新的打包格式或协议,但这应该是最后的措施。
- 样本数据:流模式测试用的样本数据的准备。
- 测试和反馈:应该包括一个原型或模拟的流式数据传输环境。

B-6 准备

启动阶段包括以下活动。除第一项活动(确定需求)外,大多数其他活动可以并行进行。

- 1) 确定数据产品可满足的需求:这通常是外部环境结果,如服务组合的定义、应用程序驱动对新类型信息的需求、航运的新法律要求和 IMO 决议等。
- 2) 定义产品的范围:主题域、预期包含的信息类型,同样也应确定不包含的信息。
- 3) 确定是否可以扩展现有产品规范:如果可以,那么这种扩展可能比开发一个新的产品规范花费更少的时间和精力。
- 4) 确定产品内的子区域:即产品规范的范围,或现有产品规范的新范围。
- 5) 定义约束:域、应用和平台等。
- 6) 收集源信息的示例:通常包括现有的数据库、官方、非官方、政府和商业出版物,特别是那些广泛使用的出版物。
- 7) 定义应用用例。
- 8) 概述数据产品启用的应用程序功能。
- 9) 定义交付模式(传输集、消息和 Web 服务等)。
- 10) 获得相应赞助组织的批准。
- 11) 组建一个项目小组。

如果产品规范的制定是在 IHO 的框架下进行的,则启动阶段还需要包括以下步骤:
在决定将新的基于 S-100 的产品规范纳入 HSSC 工作计划之前,应考虑以下因素:

- 1) 是否已将所建议的产品规范的必要性记录在案?
- 2) 是否用详细的例子证明了迫切的需求?
- 3) 产品规范是否在 HSSC 的职责范围内?
- 4) 拟议的产品规范说明是否充分说明了 IHO 的成本以及可能的立法和行政负担?

5) 是否考虑了开发时长?

为简化上述决策过程,产品规范说明书应预先提供一套信息供 HSSC 审议,如表 B-6-1 所示。

产品规范说明模板

表 B-6-1

产品规范说明		
行为	描述	注释
产品规范编号	产品规范的预期编号	在 HSSC 分配编号前,将此部分保留为空
标题	产品说明书的标题	
摘要	产品规范的简要摘要总结如下: 1. 预期用途; 2. 主要用户和次要用户; 3. 预期的功能	
产品规范范围	本规范的总体范围	应包括预期的数据来源。 例如: NIPWG 出版物……来源是 S-126 的海岸引航员/航行方向(例如,如果有人认为潮汐表应该在这里而不应在 S-104,那么它就会被列在此处)
理由	制定本产品规范的原因	
规范的互操作性	与 S-100 系列中其他产品规范的互操作性	说明产品规范的限制
S-98 的适用性	是否适用于 S-98(是/否)	在 S-100 航行系统中的 S-98 互操作规范
与其他 HSSC 工作组的合作	指定哪个工作组将参与到哪个扩展	NCWG: 图示 DQWG: 数据质量的规范信息
预算	预算需求和数字说明	是否需要 IHO 预算? 开发项目会不会由外部融资? 融资的程度如何? 应该有一个通用的指导方针来计算这一点……根据现实和证据得出合理的值
计划	对计划时间范围的说明	或者至少是要完成的步骤,因为这很难确定

B-7 开发数据模型(应用模式)

B-7.1 介绍

S-100 中定义的应用模式通常与“域模型”同义,后者用于信息建模。它是与数据产品相关的类、属性和关系的规范。

通常,产品规范只包括一个应用模式(可以分解为多个关系图)。然而,在理论上,描述不同作用域的产品规范可能需要区分不同作用域的应用模式。例如,同时包含矢量和覆盖数据的产品可能需要两个应用模式。

为了最大限度地降低复杂性,产品规范开发人员应该尽量避免定义多个应用模式。
请注意,这并不妨碍使用多个关系图来描述单个应用模式。

应用模式应该只描述要包含在数据产品中的要素、信息类型及其属性和关系。理解数据产品及其在应用程序、服务或服务组合中的角色所必需的任何其他类、约束或元素(如参与者、状态图、流等)都应与应用程序架构区分开来,并在产品规范中单独记录。

模型开发人员应该尝试限制模型元素的数量,同时仍然允许进行适当的(概念上适当、逻辑上适当、一致、正确和基于性能的)区分。

牢记在关系数据库设计中学习的数据规范化原则,但是模型开发人员还应该注意,基于 S-100 的域模型和应用模式是对象-属性-关系模型,而不是数据库设计。

B-7.2 模型开发步骤

- 1) 确定数据产品是覆盖数据还是矢量数据。覆盖数据的特征是分布在一个或多个区域上的特征值,而矢量数据的特征是局部区域(点和/或区域),这些局部区域具有边界并且在特征值中不表现出内部变化(或者在可以忽略这些内部变化的情况下)。
- 2) 确定应用域中的概念。这将涉及审查原始材料,以确定域中的重要要素和信息块,这些要素和信息块将在启动阶段定义的用例的上下文中对最终用户有用。原始材料包括启动阶段确定的样本文本,以及相关应用的文档和数据字典(如果有的话),现有或假设应用的要求、相关标准以及 IEC、IMO、IALA 等机构的通知。
- 3) 使用关键字(类别、属性和关系)在 IHO GI 注册表中搜索可重复使用的现有概念。
- 4) 仅为 IHO GI 注册系统尚不存在的概念创建新概念。这涉及更详细地检查前面提到的原始材料,以确定概念及其定义。提交建议书(提案)的过程见 S-99 标准。
- 5) 定义描述域和与数据产品相关的类和属性。如果是已经在 IHO GI 注册表或现有产品规范中定义的一类、属性和代码列表/枚举,则可以重复使用。如果没有,项目团队将需要开发和定义类和属性,包括枚举和代码表类型的值。

6) 定义类之间的关系。应定义关系,以便获取那些存在于现实世界中的数据,并建立对应用程序处理有用的链接。上述两个理由通常都适用。

示例 1:类建模结构对象和安装在结构上的类建模设备之间的结构/设备关系。

示例 2:类建模引航服务区和引航服务联系人信息间的联系人信息关系。

7) 指定适用于类、属性和关系的任何约束。例如,需要对属性进行条件编码的约束、排他关系(即当一个实例只允许参与多个可能关系中的一个时)等。结构约束通常在 UML 类图中描述,不适用于单个属性(为了减少混乱)。无论是否描述,任何限制都必须执行,(如果可能)并记录在产品规范的适当章节或制品中。

8) 准备一个或多个描述域模型的 UML 类图。S-100 模型的推荐实践方法是基于经本文件第 B-7.6 节修改的 ISO TC211 推荐实践方法。

9) 准备支持性文本,解释应用程序架构的整体结构。解释每个图的用途以及和类之间的关系。此外,解释一些不明显或应该特别注意的类的其它任何细节或其关系内容。文本不应该仅仅是类、属性和关系的列表,相反,其应该通过解释图所表示的领域现象以及和类之间的关系来阐明图中所描述的模型的用途。例如 S-122 中的示例:

一些保护区要求在某些事件(如动物攻击或污染事件)发生时,向当局提交报告。其他区域要求在进入和离开时向指定的主管部门报告。这些要求通过船舶报告类与主管部门类的关联来建模。所涉地区以所需类型的要素为模型。例如,第 6.2.1.3(S-122)条所述的海洋保护区或船只运输服务区。任何时候,报告归档的要求或限制由通知时间属性描述,并在船舶报告的文本内容属性中以文本形式提供解释(如果有的话)。文本内容属性中还描述了所需的报告格式(如果需要包括的话)。(图)显示了用于携带这些条件的模型元素,注意,为了减少混乱,这并不包括所有允许的关联或角色。

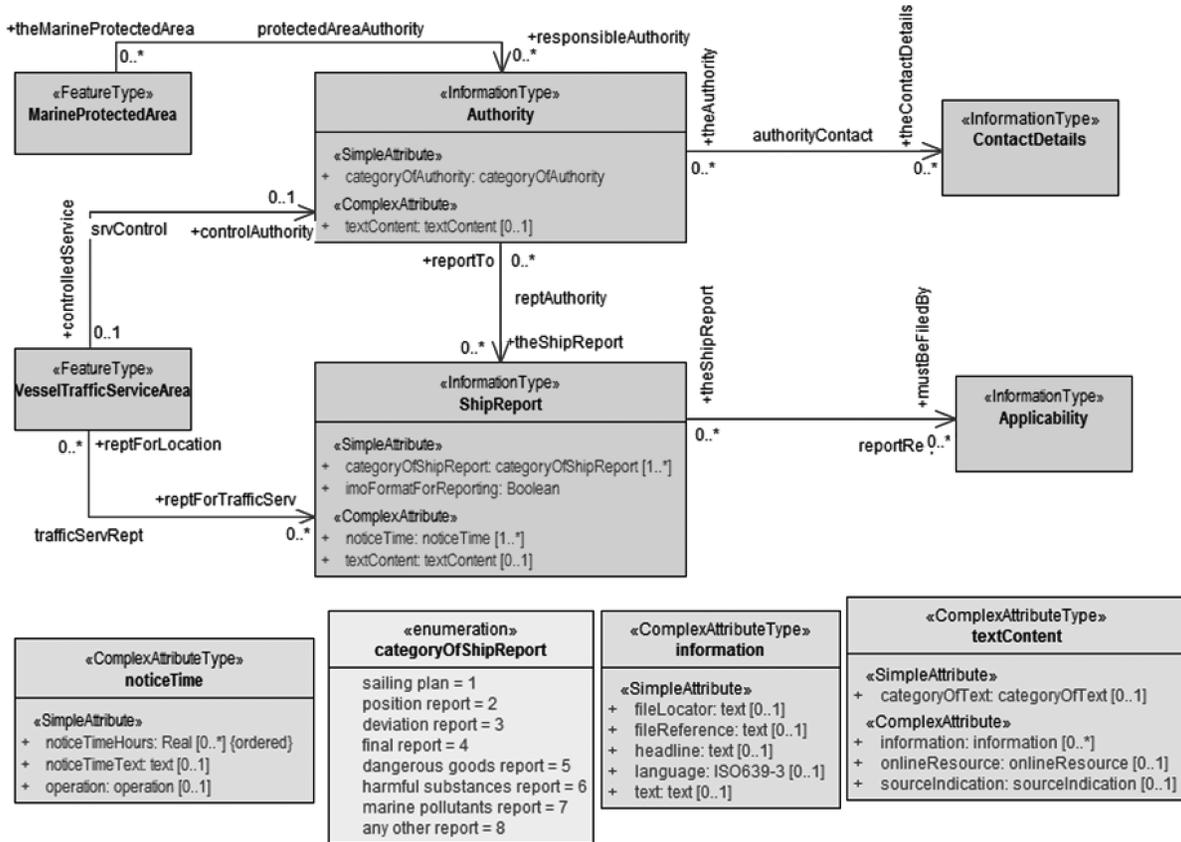


图 B-7-1 S-122 应用模式 UML 图示例(来自 S-122)

B-7.3 与通用要素模型的关系

通用要素模型是要素及其特征和关联的概念模型,是要素目录结构的基础。应用模式中的要素和信息类型必须通过 S-100 通用要素模型(GFM)的元类 S100_GF_FeatureType 和 S100_GF_InformationType 或相应元类的子类来实现(S-100 第 3 部分图 3-1)。属性必须通过 S-100 第 3 部分图 3-2 中定义的主题或空间属性元类或子类实现。产品规范可以定义派生其所有要素和信息类的本地根类,如图 B-7-2 所示,或者可以分别通过 S100_GF_FeatureType 和 S100_InformationType 实现要素和信息类型。

B-7.4 应用模式规则

应用模式的 S-100 规则基于 ISO 19109。S-100 应用模式的规则见 S-100 第 3 部分第 3-6 和 3-7 条。

B-7.4.1 矢量数据应用模式

矢量数据应用模式的规则见 S-100 第 3 部分第 3-6 条。主要规则概述如下:

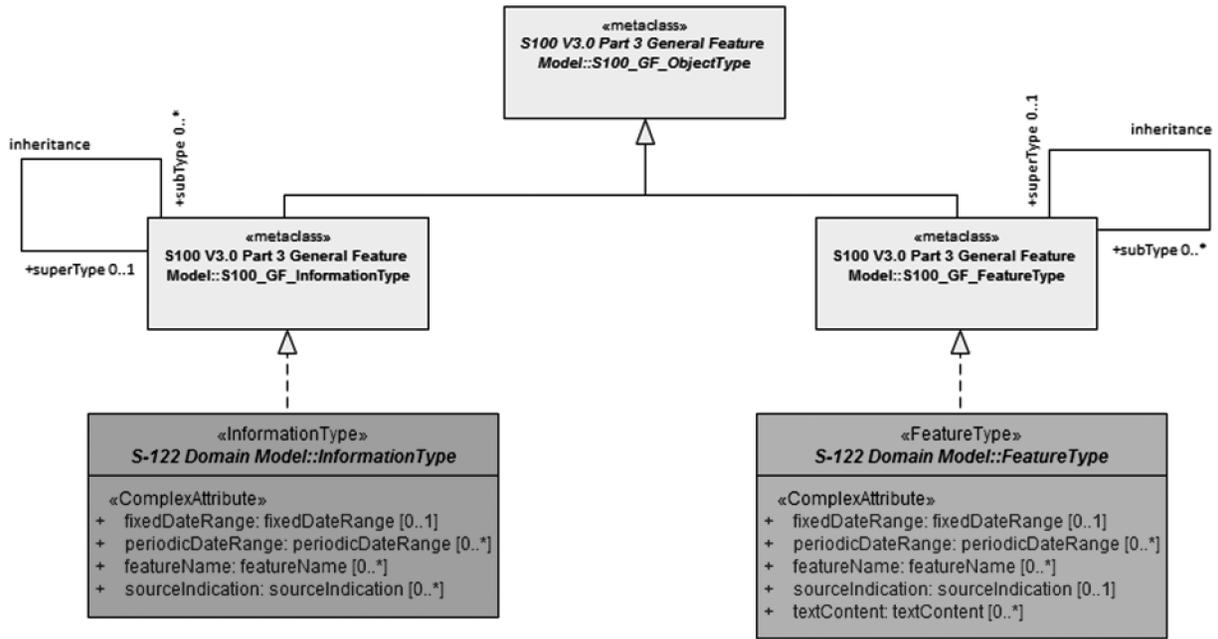


图 B-7-2 S-100 GFM 实现示例

- 要素、信息类型和复杂属性必须以类建模。
- 关系由 UML 关联或关联类建模(后者仅在关联本身具有属性时才建模。参见本文中的第 B-7.5.4.2 条)。
- 属性由 UML 属性使用适当的类进行建模。
- 关联必须被标记(具有关联名称)。还必须标记可导航的关联结尾(应具有角色名称)。

注:图可以抑制标签的绘制,从而保证清晰和减少杂乱。标签可以由产品规范文本中给出的特定规则来定义,而不是由 UML 图来定义(例如,允许产品规范使用类似“theFeatureX 在其所有关联中的角色是 FeatureX”(参见 S-100 第 3 部分第 3-5.4.5 条,关于关联结束的默认名称)的语句来“标记”关联结束。
- 空间属性必须建模为具有空间模式中允许的空间类型中的一个或多个(即联合)数据类型的属性,或者表示要素的类与空间架构中定义的空间对象之一之间的关联。
- 枚举类型及其列出的值必须由 UML 枚举建模,代码列表必须用 S-100 第 3 部分第 3-6.7 条中指定的标记建模为 UML 类。
- 标准模式(例如空间模式、要素目录模式)不应在应用程序模式内扩展。
- 应用模式中用于数据传输的所有类都应该是可实例化的。这意味着集成的类不能是构造型 <<interface>>。
- UML 应用模式应在 UML 包中进行描述,该 UML 包应带有应用模式的名称和包文档中所述的版本。

B-7.4.2 覆盖数据应用模式

S-100 第 3 部分第 3 ~7 条和第 8 部分给出了覆盖数据应用模式的规则。这些规则与矢量数据的规则相似,但有以下区别:

- 覆盖要素的空间类型由 S-100 第 8 部分中定义的点集、网格或 TIN 类型建模。

B-7.4.3 附加规则

要素和信息类型的名称必须使用驼峰命名法进行编码。矢量要素类必须使用构造型 <<FeatureType>>, 信息类必须使用构造型 <<InformationType>>。构造型允许设计人员通过创建新的模型元素来扩展 UML 模型。符合 S-100 的覆盖类型元素(描述覆盖几何)应该使用 S-100 第 8 部分中适当的构造型,覆盖数据的应用模式可以通过定义具有主题数据属性的 <<FeatureType>> 元素,并将其与覆盖类型元素相关联来描述数据属性。

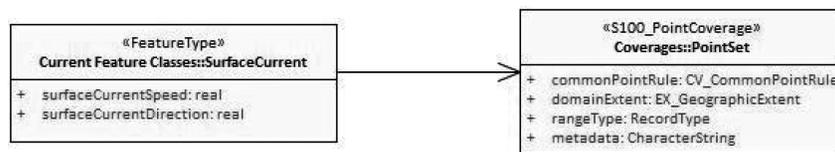


图 B-7-3 符合 S-100 的覆盖要素和类型元素示例(来自 S-111)

如有必要,产品规范除了使用标准的原型之外,还可以使用域特定的原型。

抽象类必须通过斜体化类名进行标识(如果在 UI 中选“抽象”复选框,则 Enterprise Architect 会自动执行此操作)。

S-100 指出,“应尽量减少使用多重继承,因为它往往会增加模型的复杂性”。多重继承是指一个类具有多个直接父类的情况。应用模式开发人员应该注意,多个继承与 S-100 GFM 相抵触,后者允许要素和信息类型最多具有一个父类型。

B-7.5 其他约定和建议

B-7.5.1 重用和协调

在定义新元素之前,应检查 IHO GI 注册表中是否存在可重复使用的已有元素。还应考虑以下几点:

- 要素、信息类型和属性应尽可能重复使用。
- 结构和关联应与 S-101 和其他现有的相关或补充产品相协调。
- 除非绝对必要,否则应避免定义相似却略有不同的项。

- 可以向注册系统提议扩展项(如枚举中的其他列出值),如有可能,必须避免发生同一术语的不同定义等产生的冲突。
- 通过添加特定于产品的约束(例如限制枚举或代码列表类型的允许值集),现有项可能是可重复使用的。例如,重复使用在其他产品规范中定义的复杂属性,并带有排除其某些子属性的限制。

B-7.5.2 要素和信息类型

要素类是“真实世界现象的抽象表示”(S-100 第3部分 第3-5.1.1条)。应用模式应定义一个用于描述的要素类:

- 1) 实例具有空间(地理)位置的概念。此类代表一个地理要素。
- 2) 概念的实例是上述的集合。集合可以由一个这样的类的实例或多个这样的类的实例组成。根据集合的性质,此概念可以是聚合要素或组合要素。在应用模式中,它们被视为地理要素,但可能具有空间属性,也可能不具有空间属性。
- 3) 制图信息(例如要素名称、注记、罗经圈、图例)用于定位在指定位置(或者必须从其相对于要素的默认位置重新定位),以避免模糊其他要素。这些信息虽然不是任何现实世界现象的表示,但也被建模为要素,并在要素目录中作为要素处理。这种要素称为制图要素。要显示的信息可以被编码在制图要素类的属性中,或者通过参考另一要素类来识别。
- 4) 与特定区域中的所有要素(或要素的定义子集或定义属性)相关的元信息或元数据被称为元要素。

S-100 要素目录模型提供了一个 featureUseType 元素,用于将要素类型指定为“地理”“制图”或“元”。聚合和组合要素类型被指定为“地理”featureUseType。

信息类型是可识别的对象,可以与要素或其他信息类型相关联,以便将特定信息携带到相关联的对象。信息类型通常用于在要素和其他信息类型之间共享信息。信息类型仅具有主题属性。

B-7.5.3 父类和子类

当模型中存在3个或更多概念上相似的类时,建议定义抽象父类。相似的类具有一些相同的属性或关系和允许值,共享相同属性。类不必绑定完全相同的属性集或具有完全相同的关系。

请注意,除非显式重写,否则关联和属性都是由子类继承的。虽然允许重写,但不鼓励使用它,因为它在概念上以及在数据格式和实现中引入了额外的复杂性。取而代之,应更改模型让其在层次结构中向下移动属性或关联,并只将其分配给允许使用它的类。

B-7.5.3.1 子类与类别属性

对同一物品有不同的建模方法。如,为了指示不同的浮标类型,可以使用子类方法或枚举法(category Of)。图 B-7-4 描述了这两种方法。

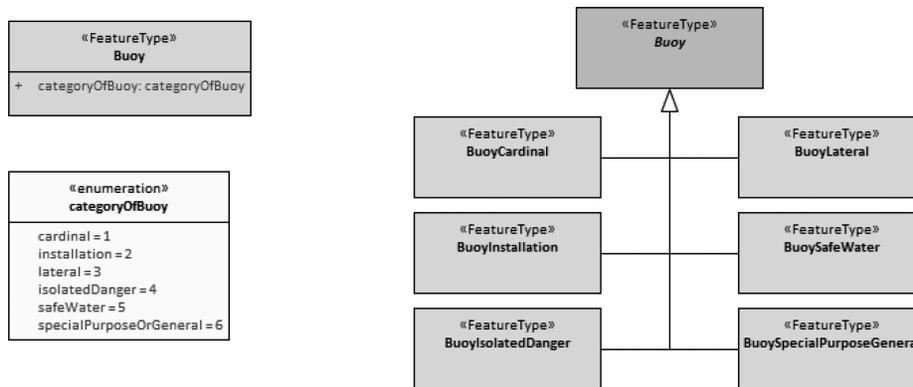


图 B-7-4 使用 category 属性和同级子类的替代模型说明

由产品规范开发团队决定哪种方法最适合他们的需求。为了确定这一点,应考虑以下问题:

- 1) 这两种方法中的任何一种会导致与外部源的显著差异吗?
- 2) “类别属性”方法是否会因为类型符号的不同而导致图示表达的问题? 或者子类方法是否会因为符号相同而导致不必要的显示规则?
- 3) 任何子类都有自己的特定属性或关系吗? 如果是,最好采用子类方法。
- 4) 不同的类别/子类是否位于(或可能位于)不同的视图组中,或具有不同的绘制顺序? 如果是,则会“稍微”倾向于创建子类。(只是“稍微”,因为图示规则和互操作目录可以使用属性值将视图组分配给要素实例)。
- 5) 如果使用子类,是否会引入对不同类别的重合对象进行编码的情况? 如果答案是“是”则更倾向于类别方法。
- 6) 子类在概念上是否非常不同? 如果答案是“是”则更倾向于子类方法。
- 7) 哪种方法可能与现有数据库和实现等外部资源兼容?
- 8) 如果有大量的子类型,那么类别方法可能更可取,因为它导致 UML 图中更紧凑的表示和更紧凑的 DCEG。(“大量”显然是主观的,但根据对人类认知心理学和用户界面中可能的实现方法的研究,一般在 5 到 9 之间,根据概念语义和相似性而有所不同)。

9)应用模式和要素目录的总体复杂性。要素(或信息类型)的同级子类比类别属性生成更多的制件和文档。这意味着要素为 DCEG 中的每个子类添加一个表,要素目录中的每个元素都有一个额外的 XML 元素,以及每个类的 UML 关系图应用模式中的一个框。从这个意义上讲,同级子类对编码器和开发人员来说是一个更大的认知负担。

B-7.5.4 关联和关联类

B-7.5.4.1 导航性、源和目标

如果关联只能在一个方向上导航,则应指示关联的导航性。也就是说,模型设计者希望应用从一个对象访问另一个对象,但反之则不然。要素/要素和信息/信息关联通常可在两个方向上导航,而要素/信息关联必须从要素端导航,不需要建模为可在另一个方向上导航。UML 将 UML 图中的导航性信息视为运行提示而不是强制需求,实现和数据格式可以以最有效的方式自由地实现导航性。

单向导航性通常将决定关联的源和目标。

关联的源和目标应该在语法和语义上与关联的名称和定义兼容。例如,对于关联“人/订阅/杂志”,源应该是人,目标为杂志。

对于要素/要素关联,两端都应命名,对于要素/信息关联,应该命名信息端,并且可以命名要素端。这是 S-100 的要求,而不是 UML 的要求。

B-7.5.4.2 关联类

关联类是向关联添加参数(特征)的一种方法,而不是在关联结束时向类中添加参数。关联类的属性描述了位于关联末端的类之间的关系。

关联类的用例基本上是“每当一个关系以一个或多个属性为特征时”。

示例 1:一组船舶受法规约束,另一组船舶免于该法规约束。船舶由信息类型描述,法规由另一个信息类型描述,它们之间通过关联关系建模,该关联具有将关系描述为包含或排除(特定法规中指定子集中的)的属性。这个关系是通过下面图 B-7-5 中的 InclusionType 关联类建模的。

示例 2:载有特定货物和尺寸的船舶必须使用指定的引航员登船点。尺寸较小的船舶建议使用引航员登船点,军舰可免于使用引航员登船点。船舶类型由信息类型描述,引航员登船点为要素类,它们之间通过具有属性的关联类关系,来阐明是否需要/建议/豁免使用特定于此类船舶的引航员登船点。此关系由下面图 B-7-5 中的 PermissionType 关联类建模。

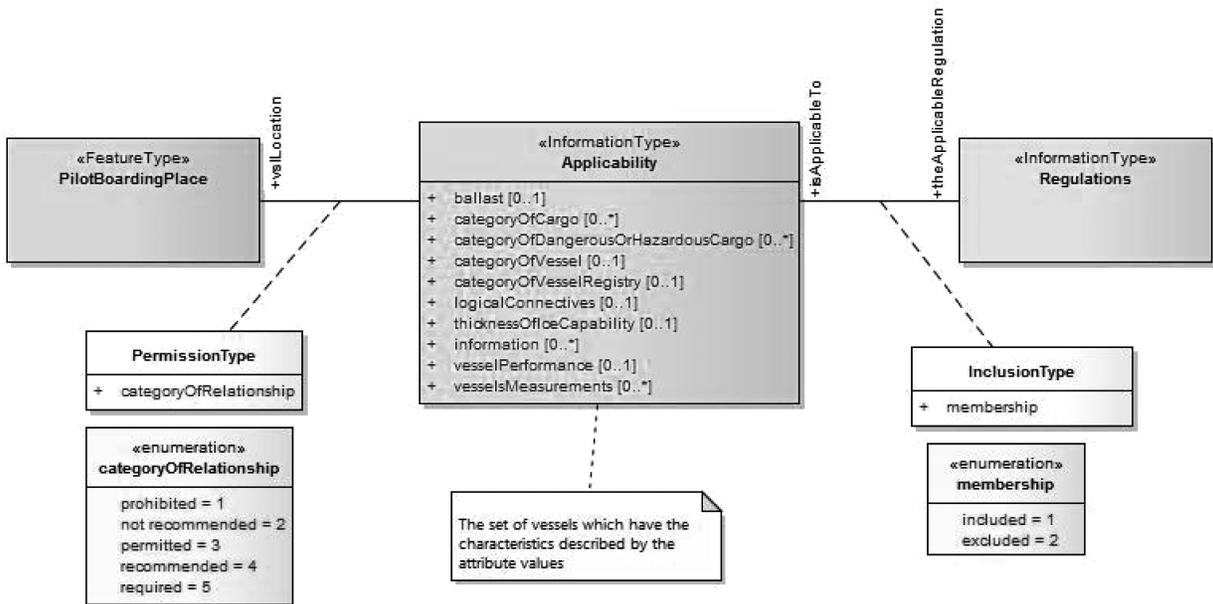


图 B-7-5 关联类建模示例

B-7.5.5 通用属性

某些属性可以使用与其他属性相同的一组列表值。例如,罗经圈点的枚举可以由属性 `windDirection-CompassPoint` 共享,它的属性是风来的方向,而 `directionOfMovement` 是描述天气系统去向的属性。目前可以简单地在属性的定义中处理。

复杂属性和空间属性可以建模为 UML 模型类元素中的命名属性,其类型对应于空间几何图元或复杂属性的名称。或者,通过关联(空间类型的普通关联,复杂属性的组合)将模型元素链接到要素/信息类。这两种方法如图 B-7-6 所示。由于附加的框和关联线,第二种方法不适用于复杂模型。

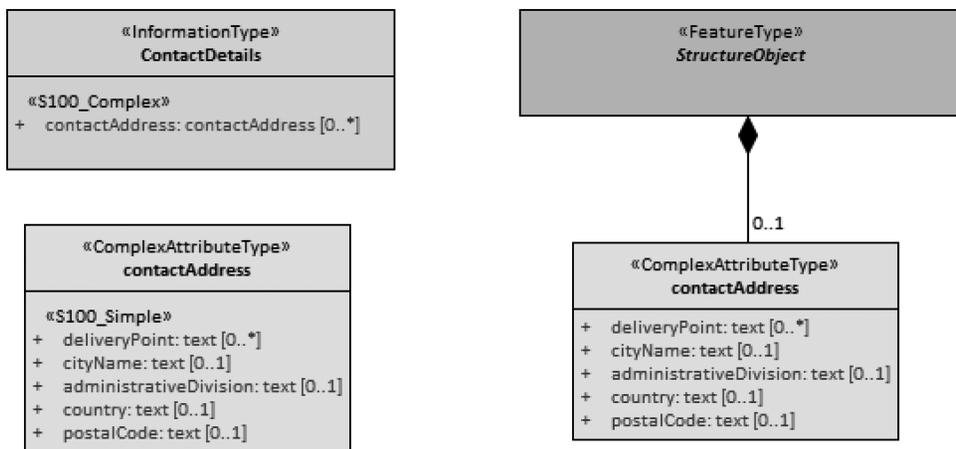


图 B-7-6 表达复杂属性绑定的方法

B-7.5.6 代码表和枚举属性

代码列表应该仅在枚举不可用或效率低下时使用(例如,如果产品规范作者不知道完整的值列表,或者允许值列表很长、不稳定、由另一个权威机构控制和/或由多个域共享)。

代码列表或单个枚举值(“列出的值”)通常仅在给定的上下文(或“容器”)中有意义,该上下文是特定的(命名的)枚举。上下文对应于属性概念。

B-7.5.7 列表值的标签和定义

标签应该简短但信息丰富,需要注意的是终端用户和编码器都会利用标签。终端用户显示和生产工具的实现者可能会使用列出的值标签作为“工具提示”或解释性文本,或者用作终端用户属性数字代码值的“显示文本”。终端用户更有可能看到标签,而不是完整的定义,这是由他们的注意力和屏幕显示等限制造成的。

B-7.5.8 数据类型

S-100 定义了一组基本数据类型和派生数据类型(第 2a 部分第 2a-4-10 条)。主题属性的属性值应该是该表中列出的类型之一。如果需要对值进行限制,产品规范可以定义限制,如果可能,可以使用 S100_CD_Constraints 中的一个或多个元素(长度、范围、模式和精度)进行编码。如果有不能这样编码的约束,则必须记录为“备注”或注。

S-100 中定义的数据类型原则上可以通过应用模式进行扩展,但如果这样做,产品规范必须根据 S-100 中预定义的数据类型定义扩展的数据类型,并在要素目录中使用预定义的数据类型,因为要素目录模式当前不允许用户定义数据类型。数据格式可以使用在底层格式标准中定义的等效内置类型(例如 HDF5 和 XML 内置类型),以便利用标准数据验证软件,前提是在产品规范或底层格式标准中记录了等效性。

B-7.5.9 列表值代码

数字代码必须是正整数,如果可能,应在 1 ~ 254 的范围内(允许数据格式和实现使用紧凑的表示,但最多允许 65535 个)。仅当提案是已作废列表值的修订版(替代)时,才能使用已作废列表值的代码。

IHO GI 注册系统还允许为列表值指定别名和驼峰编码代码。驼峰编码代码字段应使用从标签派生的驼峰编码代码来完成,以小写字母开头。驼峰编码代码准则见 S-99 附录 A。

B-7.6 推荐做法

B-7.6.1 模型元素和结构审查

在开发过程中,应经常审查模型,审查应包括相关领域专家和信息建模专家。

B-7.6.2 图的布局

应该遵循 UML 图布局的常见“最佳实践”。特别是图不应该包含太多的元素,应尽量减少线交叉,并对层次结构使用垂直布局(如果垂直布局不起作用,则使用左右水平布局)。表示关联的线条应最大限度地减少曲线段的使用。

B-7.6.3 模型元素的颜色编码

颜色编码应用于区分要素、信息类型、枚举和代码表、复杂属性、关联类以及约束和注释的图元素。抽象类型应该用深色表示。

图 B-7-1 和图 B-7-5 说明了使用颜色编码来描述不同类型的 UML 元素。将这些图中的非抽象要素和信息类的阴影与图 B-7-2 中的抽象要素和信息类进行比较。

为了区分要素和信息类型(信息类型的概念是 S-100 独有的),S-100 偏离了 ISO TC211 关于使用黑白纯 UML 图的建议。

B-7.6.4 文档表格

可以像 S-100 中的 UML 模式文档表那样进行格式化,也可以由 UML 软件生成。无论使用哪种方法,文档都必须记录类、属性、枚举和代码表类型、关联,包括名称、定义、多重性、数据类型和角色。

B-7.6.5 推荐的软件工具

S100 工作组建议使用 Enterprise Architect™ 来开发 UML 应用模式。其他 UML 工具或现成编辑器中的特殊模板也可以使用,但在 UML 符号中可能有一些小的差异,需要在产品规范中进行调整或说明。

XML 数据(包括要素目录和元数据)比普通的文本编辑器更容易在开源或 COTS XML 软件中查看。

B-7.6.6 模型标识

每个应用模式的标识应包括名称和版本。如果产品规范中只有一个应用模式,则此标识包含在产品规范的名称和版本中。具有多个应用模式的产品规范必须标识每个应用模式,可以通过将其与范围进行关联来实现。

B-8 数据分类和编码指南

数据分类和编码指南(DCEG)提供有关获取数据的信息,应尽可能详细和具体。DCEG 主要由制图员、编辑和数据编码人员使用,而不是给应用程序开发人员或原始设备制造商使用,编写时应注意角度。

DCEG 包括将现实世界对象映射到数据集的概念对象的集合标准。

执行数据获取的组织除了需要提供获取过程中使用的产品规范中所示的参考外,还应提供更详细的编码指南。

B-9 IHO GI 注册系统

S-99 标准解释了注册程序。建议至少将项目组或工作组的一名成员注册为提交组织。提交组织负责提交对注册表的内容进行修改和增补的提案。提交报告的组织通常代表一个公认的机构或利益相关方群体(如政府、工业界、学术界和相关用户群体等)。已注册的提交组织可向注册表中的任何域提交提案。利益相关方和其他有关各方应通过现有提交组织提交提案。

为了与其他产品规范协调一致,开发人员应尽可能提出对现有注册表项的扩展。例如,提议现有元素的泛化或专用化,或者枚举以及代码表类型中的附加值。现有类型的限制可以成为新的子类型,而不是更改已定义的类型,或者在产品规范中定义一个约束即可。

B-10 要素目录

要素目录是符合 S-100 XML 要素目录模式的 XML 文档。注意,对于图像和网格化数据,覆盖是一种要素类型,产品规范要素目录应该定义要素覆盖属性(具有空间图元类型“覆盖”)。

要素目录应以文本形式记录其内容,并应由项目组 and 负责的工作组进行审查。为了便于查看,应该从 XML 要素目录中生成文本文档。生成的文本可以是 Word 或 PDF 格式,也可以是团队首选的其他格式。

B-11 数据传输模式和封装

定义交付模式,即交换集、消息或服务。产品规范还可以指定多个交付模式。

该工作需要在指定元数据之前完成,因为某些元数据元素以及元数据的处理(例如,单独的和嵌入的)依赖于交付机制、约束和协议。

包装和传输内容的细节可在稍后阶段最后确定(见第 B-18 节)。

B-12 元数据

S-100 的第 4 部分规定了最低元数据要求(矢量数据的附录 4a-D,覆盖数据的附录 4b 和附录 8)。产品规范开发人员应该考虑 S-100 中列出的元数据元素是否与数据产品相关,以及哪些元数据元素适合其允许的打包和交付方法。对于相关元素,如果需要,为 S-100 附录 4a-D 或 4b 和 8 部分中列出的元数据元素定义适当的值和限制。开发者应该注意到,S-100 第 4 版中的第 4b 部分非常简单,开发团队需要使用底层的 ISO 标准和 ISO 元数据模式。

B-12.1 交换集产品元数据

本节描述交付交换集中产品的元数据。

B-12.1.1 通用元数据模型

S-100 在交换目录中为 S-100 交换集发现元数据定义了编码。图 B-12-1 是相关的类。最新批准的 S-100 发现元数据模型位于最新版本的 S-100 第 4 部分中。

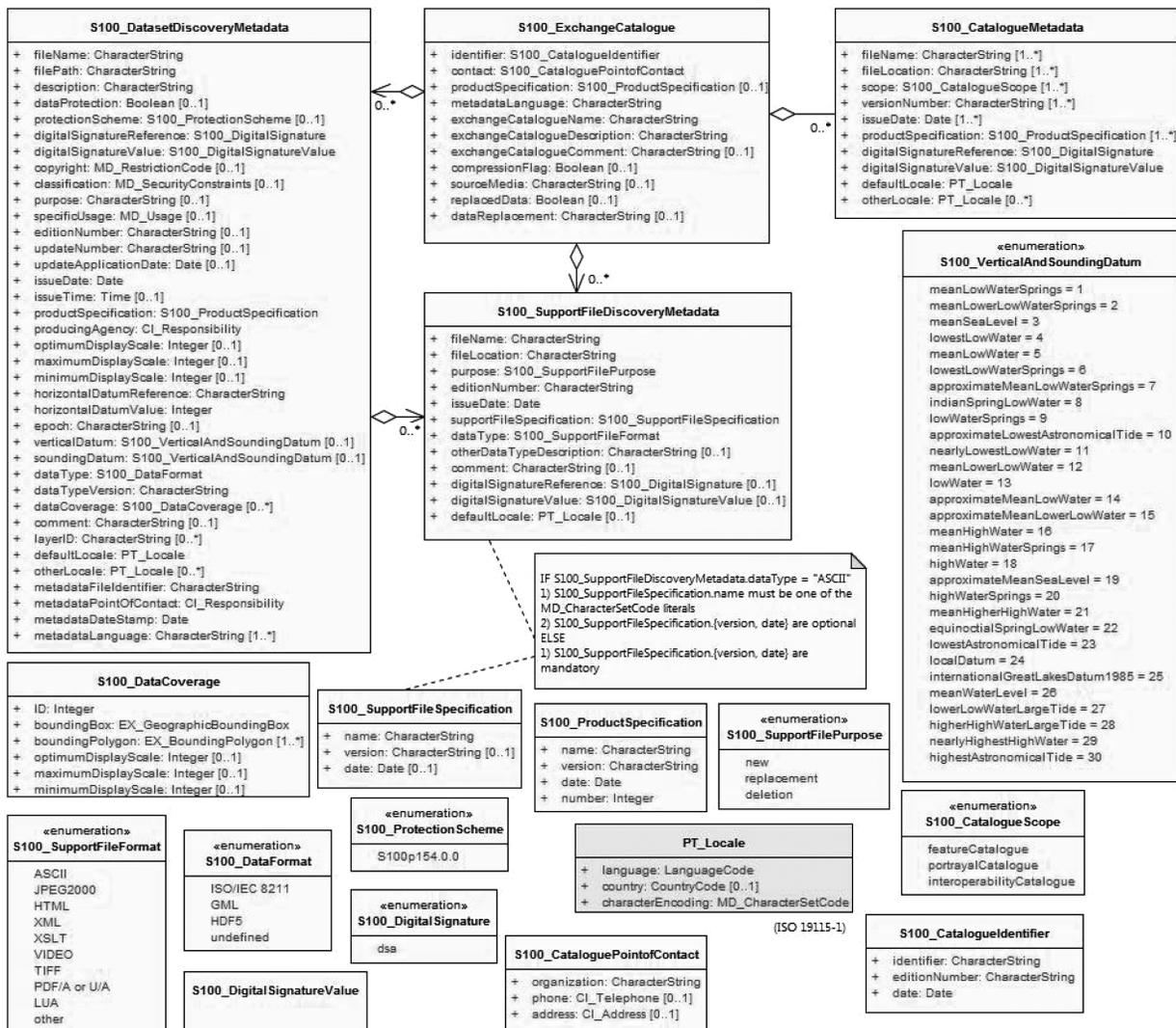


图 B-12-1 交换目录和发现元数据类

如果需要额外的元数据元素,则应在产品规范元数据部分和使用标准 ISO 扩展机制开发的通用元数据模式的扩展中记录这些元素。

在创建扩展元数据之前,必须仔细检查 ISO19115-1(和 S-100 通用元数据模型)中的现有元数据,以确认目前不存在合适的元数据。如果 ISO 19115-1(或通用 S-100 元数据模型)中存在合适的元数据,则必须使用它。(S-100 第 4a 部分,附录 4a-E,扩展到包括 S-100 通用元数据。)

IHO 交换目录和发现元数据的 XML 模式已经开发完成,可在网站 <https://github.com/IHO-S100WG> 中下载,最终会在 IHO GI 注册系统(<http://registry.ihp.int>)中呈现。

由 IHO 提供的交换目录构建器正在开发中。还可以使用现成的商业和开源 XML 编辑和创作工具来编写交换目录。通用 ISO 元数据工具可用于每个数据集的 ISO 元数据文件(在 S-100 第 4 版和 S-100 第 3 版中,ISO 元数据文件与交换目录 XML 文件分离并由其引用)。

B-12.1.2 IHO S-100 通用元数据模型和模式在交换目录中的使用

如果可以按原始形态使用通用 S-100 交换目录格式和 XML 架构,无论是否有限制,都必须使用它们。否则,产品规范必须从 S-100 第 4a 部分中的通用模型(对于网格化数据,由第 4b 部分补充)和本文第 12.1.1 条中描述的通用模型派生其元数据模型。派生可以是限制、扩展或两者的组合。派生必须符合 S-100 第 4a 部分附录 4a-E 中的规则。自定义 UML 图类似于图 4a-D-4,对于添加类或属性的产品规范是必需的,如果产品规范在枚举中省略了可选的类/属性或限制了允许值,则建议使用。

S-100 第 4a 部分指定了 S-100 元数据类和属性中哪些是必需的,哪些是可选的。产品规范可以根据需要省略可选的 S-100 元数据类或属性。

对 S-100 类和属性的限制必须表示为约束。限制包括强制可选属性或使用属性值的子集。限制也不需要新的元数据类。

可以使用从通用模型中的类派生的特定于产品的类来扩展该模型。派生类可以定义其他属性。

为了运行特定产品的元数据模型,可以补充使用通用编码 S-100 元数据的 S-100 XML 模式:

- 可以使用 Schematron³ 或其他语言来执行应用特定于产品限制的约束检查。
- 导入通用 XML 架构并扩展通用 XML 类型以获取扩展的特定于产品的架构。

³ Schematron 规则可以使用现成的软件(如 XML 编辑器)进行检查,但对于无法将 Schematron 验证或 XSLT 集成到工作流中的生产工具或应用程序,可以使用任何合适的语言来实现约束。

此方法表示不需要更改通用交换目录、XML 架构或 XSD 文件。相反,可以通过向 S-100 通用实现添加补充文件来实现特定于产品的定制。

表 B-12-1 描述了产品规范如何描述其元数据模型。这些行动是 S-100 第 4a 部分附录 4a-E 所列举的允许扩展的详细说明。UML 图将基于 S-100 图 4a-D-4 制作。

产品规范必须通过重复使用 S-100 通用元数据模型或根据 S-100 附录 4a-E 中的规则扩展通用模型来定义其元数据模型。

ISO 19115-1 和 S-100 第 4a-E 部分要求在创建扩展元数据之前,必须仔细检查 ISO 19115-1 中的现有元数据,以确认适合的元数据不存在。如果 ISO 19115-1 中存在适合的元数据,则必须使用它。

因此,应该避免对一般 S-100 元数据类进行不必要的专门化。

为各个产品规范制定交换目录中的元数据

表 B-12-1

行动	UML 图	文档表格	XML 实现
省略可选类或属性	使用绘图工具功能抑制其显示	省略相应的表格或行	用于检查 XML 元素不存在的规则
创建可选类或强制属性	无变化。多重性为 0,但表注释可以添加并声明其在产品中是强制的	“多重性”列应具有强制多重性,同时“备注”列声明其在产品中是强制的	用于检查 XML 元素存在的规则
限制多重性	对于以上内容,请进行适当修改。 这是使可选属性成为强制属性的一般情况		
限制枚举值到子集	使用绘图工具功能抑制排除值的限制	省略相应的行	用于检查该值未使用的规则
将值限制为文本、数字和字符串属性	无需改变	在备注例中指明该值	用于检查该值是指定值的规则
添加一个新的元数据类或属性	专门化 S-100 元数据类,添加新类和属性	添加新的文档表格或扩展描述原类的表格。继承属性应与直接属性进行区分	附加的 XSD 引入了通用元数据 XML 模式并扩展了它的类型

续上表

行动	UML 图	文档表格	XML 实现
在数据交换目录中添加新元数据包 (注:新包是否必须在交换目录或 ISO 元数据中取决于产品的具体考虑因素。比如它在应用程序中的预期用途)	为新元数据包添加类,并聚合到交换目录容器类(S100_交换目录)的专门化中	新类和扩展的目录容器类的新表格	如 ISO 19115-3 和 S-100 条款 4a-5.6.5) 规定。如果交换目录中需要,则把附加 XSD 导入通用元数据 XSD 并扩展它的类型
只在 ISO 元数据中添加新的元数据包	图扩展了 S-100 图 4a-1 并指明了新类和属性	新类的新表格	描述 ISO 19115-3 元数据扩展格式的扩展,同时使用新的类型扩展 ISO 元数据 XSDs
向枚举中添加值 (可以考虑通过 S-100 维护提案向 S-100 通用元数据中添加,而不是扩展产品)	专门化相关 S-100 元数据类型,并重写受影响的属性	在合适的表格中描述该重写和新数据类型	XML 模式在继承的 XSD 中重写。忽略由于通用约束检查规则引起的所有警告。如果必要的话,定义新的约束检查规则
创建新的元数据代码表元素 (扩展代码表)	在图中扩展代码表 (有一些代码表,如语言,可能对于 UML 图来说过长。)	在代码表表格中添加代码表值,或记录新值	在合适的代码表和 GML 字典文件中添加新值
创建一个新的元数据代码表,取代具有“自由文本”列表域值的现有元数据元素域 (注:在先前的行中限制字符串是一种比较简单的方案)	专门化相关 S-100 元数据类型,并重写受影响的属性。添加记录了代码表的 UMO 元素	在合适的表格中记录专门化和新的代码表	XML 模式在继承的 XSD 中重写。忽略由于通用约束检查规则引起的所有警告。如果必要的话,定义新的约束检查规则 在合适的代码表和 GML 字典文件中添加新值
强制属性没有值	无	指明强制属性为空或添加默认值	用于检查该值的规则

产品规范不应克隆和重命名通用模型中定义的类、属性和枚举,应根据需要扩展通用模型。随着 S-100 和产品规范在后续版本中的发展,此要求有助于保持特定产品的 S-100 元数据和通用 S-100 元数据的一致性。克隆和重命名类和属性将导致元数据形成不同分支。随着 S-100 和产品规范的增加,元数据分支的差异将增大,导致版本过于复杂。结果之一是需要具有独立的特定产品的元数据处理模块。对于由同一应用程序(如通用浏览器、ECDIS 或 ECS)处理的产品,这种情况是不可取的。

典型的产品规范交换目录 UML 图如图 B-12-2 所示。与图 B-12-1 相比,可以注意到数据集发现元数据中省略了垂直和深度基准的属性,以及对支持文件格式和数据格式枚举的限制。

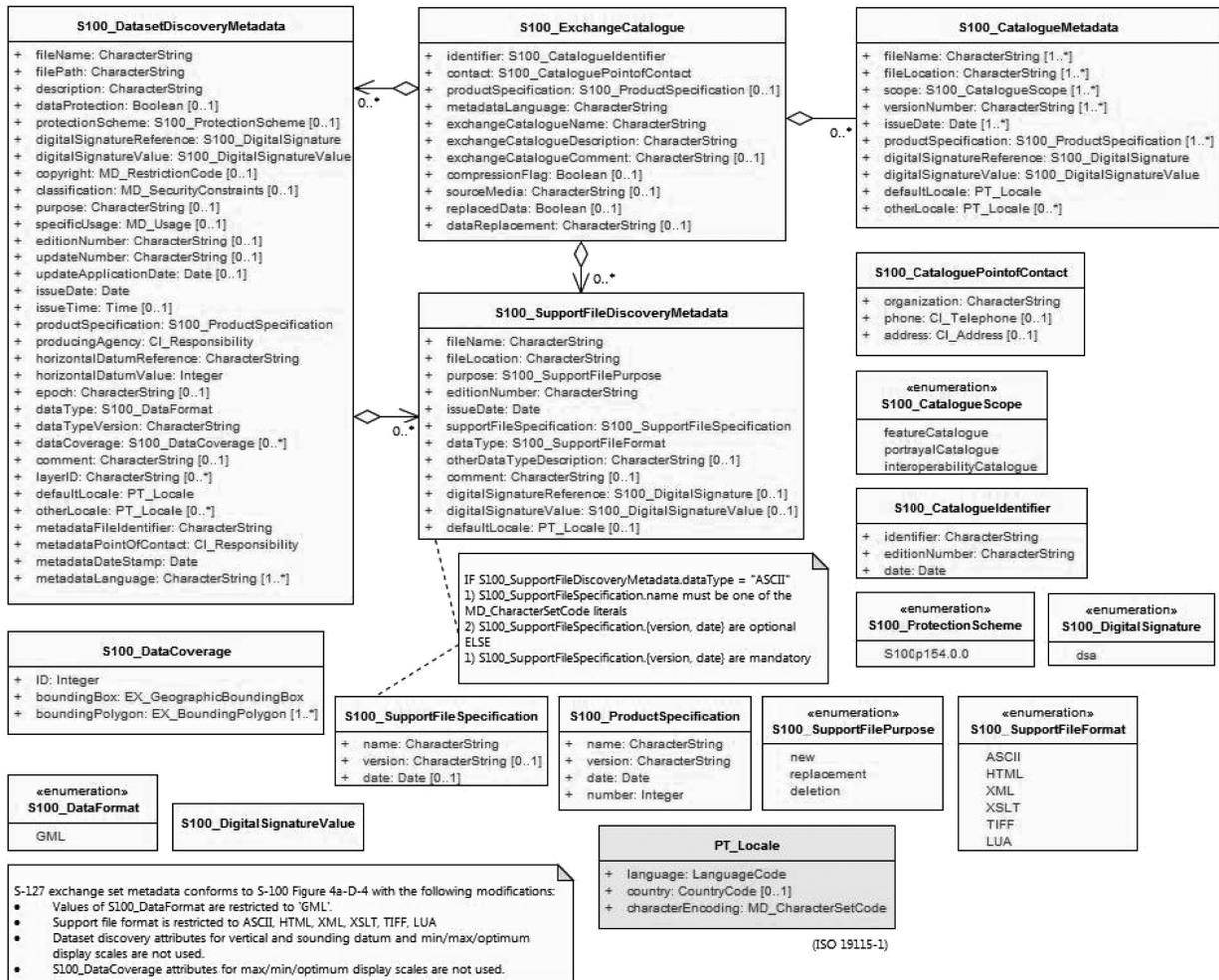


图 B-12-2 S-127 中交换目录元数据类(海上交通管理)

B-12.1.3 ISO 元数据文件中的扩展

S-100 第 4a 部分,第 4a-5.6.5 条规定了扩展必须如何被记录。该规范源自 ISO 规范。

B-12.1.4 数据保护、认证和加密

S-100 数据保护的目如下:

- 1) 防盗版保护:通过加密产品信息来防止未经授权使用数据。
- 2) 选择性访问:限制客户只能对已获得许可的产品进行访问。
- 3) 认证:保证产品是由在产品包或数据集中被标识为产品生产者和销售者实际创建和分发的。

防盗版保护和选择性访问是通过对产品进行加密并提供解密产品的数据许可来实现的。数据许可有一个到期时间,以便客户在许可期限内访问产品。认证是通过应用于产品文件的数字签名来提供的。通过向用户提供一套经许可的数据许可证,支持有选择地访问个别产品。许可证使用目标系统的唯一硬件标识符创建,并且对于每个数据客户端都是唯一的。许可证不能在各个数据客户端之间交换。

B-12.1.4.1 保护措施的应用

产品规范应使用 S-100 推荐的安全方案,使用 S100_DataSetDiscoveryMetadata 的 dataProtection 属性确定数据集必须、不能和可以进行加密。如果它的布尔属性设置为 TRUE,则还必须从安全方案枚举中为共同属性 protectionScheme 分配一个值(S100_ProtectionScheme)。S-100 目前仅定义了一个推荐的安全方案,该方案在第 15 部分中进行了描述。

B-12.1.4.2 数字签名

S-100 第 4a 部分规定交换集中的数据集和目录必须使用数字签名。对于支持删除文件,允许使用数字签名,但可选。因此,产品规范作者应该指定哪些支持文件必须、可以或不可以签名。签名数据集与是否将推荐的安全方案应用于数据集无关。

数字签名方法的编码在属性 digitalSignatureReference 中。还有一个对数字签名本身进行编码的属性 digitalSignature。这些属性存在于数据集发现元数据,目录元数据和支持文件发现元数据(S100_DatasetDiscoveryMetadata、S100_ExchangeCatalogue 和 S100_SupportFileDiscoveryMetadata 类)。关于数字签名结构规定在 S-100 第 15 部分。S-100 4.0.0 第 4a 部分在格式方面规定非常宽松,允许使用格式良好的 XML(来自未指定模式的一个或多个 XML 元素)或字符串(例如 base64 编码的签名)。

B-12.1.4.3 压缩和存档格式

S-100 第 4a 和 15 部分中使用的数据产品的压缩包括存档格式的规范以及实际的文件压缩方法。在 S-100 4.0.0 版中只有一种存档格式 (ZIP) 和一种允许的压缩方法 (DEFLATE)。压缩需要打包到存档中。产品规范作者必须指定交换集必须、绝不可以或可以通过在交换目录中的 `compressionFlag` 元数据属性上指定适当的约束来压缩。有关存档格式和压缩方法的更多信息,请参阅 S-100 第 15 部分第 15.5 条。S-100 4.0.0 版为交换目录提供了每个 `compressionFlag` 属性只有一个实例,因此该属性适用于交换集中的所有文件(即,在压缩之后,在交换集中将只有一个包含所有数据文件、支持文件和目录的 ZIP 存档,将 DEFLATE 压缩方法应用于所有文件⁴)。

产品规范作者应注意,一个交换集可以包含其他交换集,并且每个包含的交换集都可以被视为用于压缩目的的单个交换集(即,是否打包到 ZIP 存档中)。但它们都将被打包到整个容器交换集的存档中,要么作为文件夹层次结构,要么作为单个文件 ZIP 存档,这取决于它们是否被单独压缩。

不使用 ZIP 的加密和数字签名功能。

B-12.1.5 图像和网格数据的元数据

处理图像和网格信息的产品规范的交换集元数据与普通矢量数据集(条款 B-12.1.1 至 B-12.1.4)相同,扩展了额外的元数据元素和特定于图像或网格数据的属性。S-100 第 4b 部分描述了 ISO 19115-2 中详细定义的附加元数据元素。

这类产品的交换目录与矢量数据产品的交换目录相同。对于矢量数据,产品规范编制组可以根据 S-100 附录 4a-E 定义其他特定于产品的元数据元素和属性。ISO 定义的元素和属性应该在 ISO 元数据文件中。

B-12.1.6 嵌入式元数据

某些元数据可以嵌入数据集标题(ISO 8211 和 GML 格式 S-100 部分 10a 和 10b)或定义属性和组(HDF5 格式第 10c 部分)中。虽然指定嵌入的元数据符合单个产品规范,但本指南建议只应嵌入被认为对识别和读取数据集至关重要的元数据。例如数据集名称、持久全局标识符和 MRN(如果有)、边界、每种类型的对象数和 CRS 标识。网格化数据产品可以将网格间距和网格边界框等空间表示信息编码为单个要素的元数据。网格化数据产品还可以编码结构化元数据,例如表示网格类型(规则、不规则等)的代码。

4 一些 ZIP 实现可以检查压缩方法是否实际生成了一个小文件,如果没有,则存储原始文件。即使文件大小不减,也有可能进行了压缩。

B-12.2 服务元数据

本节介绍作为服务交付产品的元数据。

B-12.2.1 服务的通用元数据

服务的通用元数据正在开发中。S-100 第 4a 部分定义了符合 ISO 19115-1/2 模型的服务元数据模型。S-100 第 14 部分以不同的方式定义了服务元数据(S100_ServiceMetadata 第 14-8.1.1 条)。由产品规范来确定服务器是否需要按照底层 ISO 标准提供服务标识元数据(参见 S-100 图 4a-A-2)。

需要定义服务元数据(标识和描述服务本身)的产品规范应该使用或扩展 S-100 图 4a-2 中的 SV_ServiceIdentification 类。任何扩展或限制都应符合 S-100 第 4a 部分附录 4a-E 中所述的 ISO 规则。

B-12.2.2 IHO S-100 元数据模型和模式的使用

S-100 第 14 部分第 14-9.1.3 条规定了一个 GetMetadata() 操作,其字符串参数 exchangeSet 被定义为“描述数据集的交换集”。目前尚不清楚这是 S-100 第 14 部分第 14-8 条中定义的元数据还是 S-100 第 4a 部分图 4a-D-4 定义的元数据。可以将其视为有效负载元数据(描述传输信息的元数据),将其与描述服务本身的元数据区分开来。

如果产品规范需要定义 S-100 交换目录的子集和/或扩展以描述有效载荷元数据,则应使用 S-100 第 4a/4b 部分和本指南第 B-12.1 条中规定的相同方法。请注意,这些章节要求符合 S-100 第 4a 部分附录 4a-E 中描述的 ISO 规则。

B-12.2.3 数据保护,认证和加密

S-100 第 15 部分的许多数据保护、认证和加密规定也应适用于服务,但细节和实施限制将有所不同。例如,数据产品可以使用 ZIP 以外的格式创建其传输包,并且连接协议可能需要安全措施(例如使用 HTTPS 而不是 HTTP)。鼓励需要在服务传输模式中使用安全方案的产品规范团队探索使用或改编 S-100 第 15 部分并报告其经验。

B-12.2.4 嵌入式元数据

第 B-12.1.6 条中描述的嵌入式元数据具有类似考虑。传输模式可能引起其他考虑,例如,更加重视最大限度地减少数据量。鼓励产品规范团队考虑使用或调整第 B-12.1.6 条中的嵌入式元数据原则以及所使用的数据格式(ISO 8211、GML、HDF5 或其他格式)的方法和约束,并报告他们的经验。

B-13 定义数据编码格式

B-13.1 编码格式选择

编码格式应根据产品类型和其他要求选择,包括生产和加工等的要求。方便起见,下表总结了 S-100 4.0.0 版所包含的三种标准数据格式的特点。

S-100 标准编码格式特点

表 B-13-1

	ISO 8211	GML	HDF5
适合的产品类型	航海图和要素为核心的矢量数据	航海出版物和信息为核心的矢量数据;离散天气信息;例如海上安全信息的小数据集;通过消息和 Web 服务交付的数据	基于覆盖的数据
通用数据格式	是	是	是
数据生产复杂性	需要定制工具	可以使用文本编辑器,也可以使用定制的应用程序和数据库查询	定制应用程序,使用现有的库
处理复杂性	高	低	高
支持现有软件	是	现有浏览器和服务器软件,可以使用二进制文本编辑器查看	现有软件
数据体积	小	大	小
数据类型	矢量	矢量、GML 规范中定义的覆盖模式(S-100 中不适用)	网格
产品规范需要的支持制件	要素目录	数据验证用的 XML 模式;可以使用应用程序处理的数据集(无需 XML 模式);自归档格式(标签指示对象和属性);要素目录可选	嵌入的对象和属性标签;要素目录可选

虽然除了三种标准编码之外还可以使用其他格式,但使用非标准格式会产生以下影响:

- 失去通用性,需要实施者进行特殊目的的开发或转换为对性能有影响的标准格式。

- 可能损失互操作性及兼容性。

对于事务性、web 服务或基于消息的信息的适当格式问题,目前还没有正式深入解决,但是需要这种交付模式的产品规范应尽量使用其中一种标准格式,以便最大限度地降低实现的复杂性。

B-13.2 GML 数据格式定义

选择 GML 格式需要将封装 S-100 应用模式的 XSD 文件定义为符合 GML 规范的 XSD 文件(ISO19136 和 S-100 第 10b 部分)。在将模型实现为 GML 应用程序模式的 GML XSD 文件中,应用程序模式与 XML 模式结构之间必须 1/1 映射。

- UML 域模型中的要素和信息类映射到 XML 复杂类型和 XML 元素。类(驼峰编码)名称与元素名称相同(XML 标记不包括任何命名空间)。
- UML 域模型中的简单属性映射到包含在上面提到的要素和信息复杂类型中的 XML 元素。名称(XML 标记,不包括任何名称空间)与 UML 属性的名称(驼峰编码)相同。
- UML 域模型中的复杂属性映射到第一个项目中提到的要素和信息复杂类型中包含的 XML 元素。名称(XML 标记,不包括任何名称空间)与 UML 属性的名称(驼峰编码)相同。这些 XML 元素的类型具有相同名称(驼峰编码)的 XML 复杂类型,可以带或不带“Type”后缀。
- 枚举映射到同名的 XML 枚举(驼峰编码)。
- 代码表映射到同名的 XML 联合类型(驼峰编码),联合类型的一个成员是所列值的匿名枚举,另一个成员是一个字符串,限制为模式“other:...”,如 S-100 第 3 部分第 3-6.7 条所规定。
- 关联被编码为包含在要素和信息复杂类型中的元素,根据 UML 域模型中指示的角色命名,并使用 GML 属性包含被引用对象的 gml:id。

模式开发人员应该注意,与 GML 规范的一致性需要遵循 GML 规范(ISO 19136/OGC 07-036)中规定的规则,而不仅仅是针对 GML 模式的 XML 验证。

S-100 第 10b 部分描述了 S-100 GML 配置文件。S-100 GML 配置文件的 XSD 文件可以在 IHO 网站(<https://github.com/IHO-S100WG>)中下载。早前其他产品规范定义的 GML 模式(例如 S-122、S-123)可以提供有帮助的指南,可以在 IHO 产品规范分发页面找到下载,或在 GitHub 中下载。

GML 数据集通用处理指南包含在 S-100 4.0.0 版中。

B-14 图示表达元素和规则

计划以图形方式显示的数据产品(而不是文本或其他处理)则需要图示表达目录。

通常只有一个预期的应用领域,就是 ECDIS 及相关使用场景(航线监测/航线规划)会使用图示表达目录。有些数据产品可能不用于显示,在这种情况下,不需要编写图示表达目录。有的数据产品可能用于多个应用领域,在这种情况下,应该考虑需要不同的图示表达目录。不同的应用场景可能使用不同的图示表达目录。

S-100 在第 9 部分中定义了一个基于 XSLT 的图示表达机制。S-100 4.0.0 版添加了 LUA 脚本语言作为第二种图示机制,并定义了基于 S-100 的产品中包含脚本支持的标准机制。开发人员/OEMs 可能更喜欢 LUA,因为它可以保证性能并与公共编程语言相似。

产品规范的图示表达部分应包括:

- 符号和颜色的图形表示,并附有符号规范(后者最好是以机器可处理的文件形式以及图示部分文本中的形式规范)。
- 实施者使用符号的规范和建议,例如,根据属性值计算方向,通过数据综合减少小比例尺下的图面拥挤,显示伴随符号的文本,掩蔽边界等。
- 需要包含在 ECDIS 中的产品必须定义白天、夜晚和黄昏调色板。

B-15 图示表达元素的注册

图示表达元素必须在 GI 注册系统图示表达注册表中注册。原则上,图示表达元素只有在注册表中是新的⁵时才需要注册,已经在注册表中的元素不需要再次注册。需要注册的元素如下所示。

- 新符号、字体、线条样式、区域填充和像素映射。
- 新显示模式。
- 新视图组层。
- 视图组。
- 新显示平面。
- 新颜色标记。
- 本产品图示表达使用的新颜色配置文件。如果有多个调色板(例如白天/黑夜/黄昏),则颜色配置文件必须定义所有调色板。
- 新的 CSS 文件,对本产品图示中所有符号使用的颜色标记和 RGB 值进行编码。
- 新的上下文参数。

⁵ 这显然在不同的数据产品之间引入了依赖性。需要制定管理这种依赖关系的图示注册表设计和业务流程。在此之后,本节会在该工作完成后进行修订。

- 新的显示优先级。

B-16 图示表达目录

产品规范开发人员可以在 IHO GI 注册系统的帮助和指导下,或者在存储库部分获得 IHO 图示目录生成器。

B-17 参考系统

首选的坐标参考系统是基于 WGS 84 水平基准的 EPSG 4326。

水平基准通常通过在 EPSG 注册表中给出其代码参考。如果坐标参考系统不是 EPSG 注册表中的坐标参考系统之一,且数据产品范围内存在用例,则可以在支持文件中指定数据,如 S-100 第 6 部分所述。

S-100 附录 4a-D(S100_VerticalAndSoundingDatum)中列出了一组垂直基准。如有可能,鼓励产品规范开发者采用 S-101 ENC 基准(标准或本地 S-101 ENC 基准)作为共同的垂直参考基准。如果确定需要增加一个基准,则应提议将其作为 S-100 的订正。

S-100 包含“本地基准”作为基准属性的允许值,但即使在数据产品中,这也是有限的。[S-101 举了一个在非潮汐盆地中使用本地基准的例子,其中深度可以指不同于开阔水域中的深度基准。如果该区域在 ENC 数据的最大显示尺度上是可航行的,则必须使用属性垂直基准 = 24(本地基准)来编码该基准的值,在元要素中与码头覆盖区域重合。]将数据转换为本地数据或从未指定的本地数据进行转换都是有问题的。

B-18 数据产品交付

B-18.1 交付内容和结构

定义交付包的内容和结构:交换集、消息或 Web 服务。

B-18.1.1 交换集

应按原样使用 S-100 附录 4a-D(转载于下文)中的结构图来定义交换集结构,或限制允许的组件,或定义单个组件的扩展。如果交换集中存在内部结构(例如文件夹和子文件夹),请确定所需的布局 and 命名约定。确定整个交换集的打包方式(例如,普通文件夹、zip 文件等)。

确定哪些命名约定(如果有)必须应用于交换集的各个组件-数据集文件、交换集和支持文件。对于基于消息或服务的数据传递模式,通常不需要命名约定,但是对于每个消息或服务事务,通常需要唯一的标识符。

B-18.1.2 服务

S-100 第 14 部分描述了三种类型的通信:

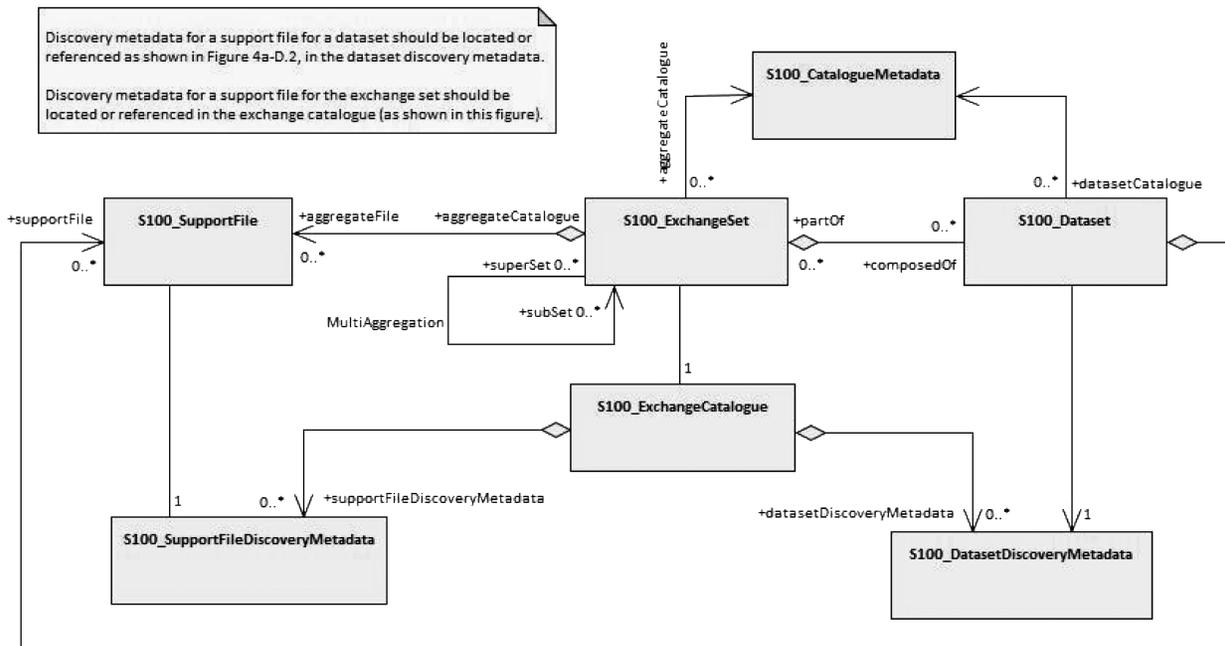


图 B-18-1 原型交换集结构

- 应用于会话的通信(S-100 第 14 部分第 14-4 条):客户端和服务端之间的点对点连接,由描述会话生命周期(会话的启动、维护和终止)的交互模型描述。
- 无会话交互通信(S-100 第 14 部分第 14-5 条):所有请求中的相关信息,服务器根据这些信息进行响应。
- 信息流(S-100 第 14 部分第 14-6 条):包含定义良好的信息的单向流数据集。

对于消息和 web 服务的传递模式,请指定容器格式、打包,并指定传输协议[例如 REST(表示状态传输)、SOAP(简单对象访问协议)]和打包[例如 WSDL(网络服务描述语言)、WFS(网络要素服务)]。S-100 第 14 部分目前允许三种服务技术:SOAP、REST 和 CORBA。

(在线服务的开发规范将包含在本文档的更高版本中,更多关于 S-100 第 14 部分的经验待定。)

B-18.2 数据集更新

B-18.2.1 更新的一般注意事项

定义数据更新的条件和机制,特别是:

- 更新周期:必须更新数据的周期,是定期发布、视需要发布还是两者兼而有之;
- 说明每个数据集的有效期和有效期的标识;
- 数据集的取消、替换和重新发布机制;

- 用于更新的元数据;
- 更新类型:数据产品是需要增量更新、全数据集替换、不规则更新或累积更新;
- 确定必须用新数据集替换的时间,以及取代、重新发布、更新和取消数据集的标准。

B-18.2.2 特定格式的更新注意事项

ISO 8211:S-100 第 10a 部分描述了 ISO 8211 数据集的更新机制。每个产品规范都必须定义更新数据集的结构。

GML:建议使用替换整个对象的方法,也有允许更新单个属性的 XML 规范。(详情将载于本文件后续版本,待进一步讨论。)

HDF5:要素(覆盖范围)可以全部或部分更新(可以是子网格的更新)。

B-18.3 支持信息

描述辅助内容如何随数据或作为数据的附件进行交付。S-100 提供了需要包括在交换集中的“支持文件”。支持文件可以由数据集对象引用的图形或文本信息文件,也可以是其他文件,如字典和目录(包括要素目录或图示表达目录)。为支持文件定义允许的文件格式和命名约定。

请注意,由于符合产品规范特定版本的所有数据集共享要素和图示表达目录,因此,相较于每个交换集一起交付,一次交付要素和图示表达目录会更有效。这种特殊交付的可能方法尚未标准化,而是留给产品规范开发人员,可能包括通过一般渠道分发的特殊交换集或诸如 web 服务出版物。如果采用集中分配机制,则必须考虑到无法或很少进入分发中心的用户(例如,低带宽或没有因特网接入的用户)。

B-19 有效性检查和数据质量

至少需要两种类型的有效性检查:

- 针对单个数据集的数据集有效性检查。这些检查将数据集中的单个对象和单个数据集作为一个整体进行操作。应检查数据集中单个对象(空间、要素和信息类型)的完整性;数据集中对象之间的关联;数据集中嵌入的任何元数据或标题信息和支持数据集中引用的文件。
- 打包有效性检查用于验证包(例如交换集)的结构和内容以及附带的元数据。

目前正在开发一套通用的有效性检查(见 S-97 C 部分)。根据从有效性检查中得出的数据集统计数据,制定一套建议的数据质量衡量标准。S-97 C 部分也介绍了建议采取的措施。

B-19.1 数据集的验证检查

对数据集的有效性检查应包括：

- 1) 完整性:包括属性的填充和所需信息是否存在,无子属性的复杂属性等。
- 2) 逻辑一致性:例如缺少关联目标。
- 3) 空间一致性:例如非交叉外部边界的拓扑完整性检查,线状要素节点过多等。
- 4) 定位精度:例如报告的坐标值与接受的或已知的绝对或相对坐标值的接近程度。
- 5) 时间精度:例如报告的时间测量值与被接受为真或已知为真的已接受或已知值的接近程度,事件顺序的正确性或数据相对于时间的有效性。
- 6) 专题准确性:如与任何其他相关属性一致,并在允许范围或集合内的属性值。
- 7) 对支持文件的引用。
- 8) 产品特有的其他要求:例如加密、签名等。

其中一些问题将在目前正在开发的一套通用有效性检查中加以解决。S-97 第 C 部分描述了推荐的通用有效性检查集。产品规范开发人员应使用适用于特定产品的附加检查来补充通用有效性检查集。

B-19.2 包的有效性检查

包的有效性检查应包括：

1) 包的完整性:是否包括所有必需的组件,包括数据集、支持文件、元数据和适当的目录(例如交换集目录、要素目录和图示表达目录)。请注意,产品规范必须指明目录适合的交付方法,例如,基于消息的传递方法可能不包括交付包中的目录。

2) 包容器格式和结构:包装是否采用批准的容器格式(例如 ISO 8211、TIFF 等),以及是否在容器级别应用了适当的加密和签名。包验证检查的示例如下：

- 假设产品规范将交付指定为 zip 文件,那么容器是否为适当类型的 zip 文件?
- 如果包被安排在目录(文件夹)结构中,那么目录(文件夹)的结构和名称是否符合产品规范的要求?

包有效性检查是验证交付的包所必需的流程,如果预计将超出 S-97 第 C 部分的范围,那么产品规范编制组可能必须指定自己的包。交换集和服务交付模式的测试显然会有所不同,但是应该针对产品规范中涵盖的所有交付模式来解决验证交付包或流的问题。

B-19.3 通用有效性检查

由于某些要素、信息类型和应用模式构造用于多个产品中,因此将采取与现有产品规范相同的有效性检查,并且应参考任何与此类相关的产品规范来进行检查。特别是空间一致性检查,以及与元要素相关的一致性检查,可以预期与几个数据规范相同。

有效性检查中使用的空间操作必须是 IHO ENC 有效性检查(S-58 6.0.0 版或其后续版本)中定义的操作。

B-19.4 基本数据集与更新数据集的有效性检查

如果产品规范定义了更新数据集格式,则应检查为新的(基本)数据集开发的验证检查是否适用于更新数据集格式。

B-20 互操作性准备

如果数据产品是包含在互操作性目录(S-98)中的候选产品,那么产品规范开发人员应该执行以下活动。关于这些活动的进一步详情将在 S-100 的未来版本和国际海道测量组织出版物 S-98(正在准备中)中提供。

- 确定数据产品是否对互操作性目录中的任何产品组进行了补充或增强。
- 确定 IHO 互操作性目录是否以及如何受到新产品的影响,包括显示优先级、交错、预定义组合以及其他互操作性规则和操作的更新。
- 根据 IHO 互操作性团队的建议修订图示表达目录。
- 将应用模式与相关的产品规范进行比较,并向互操作性目录的开发人员建议与其他产品规范中的要素的相似性。这种比较应该包括要素概念相似度、属性绑定、属性值域、要素几何中的潜在差异等。
- 检查规范范围和影响信息质量的因素,这些信息也可能存在于其他产品中。例如,信息是针对一个产品中的上下文还是背景,在另一个产品中是否是重要部分,这些分析的结果应传达给互操作性目录开发人员。
- 考虑由于制图对象而产生的交互。例如互操作性模式中的文本放置问题。产品规范团队应就如何缓解此类问题与原始设备制造商进行磋商。
- 就要素是否是另一产品中要素的补充信息向互操作性目录开发人员提供建议。这些要素可以组合作为互操作性的一部分。
- 尝试协调数据集使用的最大和最小显示比例尺的效果,以避免出现一个数据产品在显示比例尺内,而补充产品不在显示比例尺内的情况。

B-21 样本数据/测试数据集

应创建足够数量的测试数据,以验证应用模式的主要特性。具体地说,第一个测试数据集应该包含:

- 每种要素和信息类型至少一个实例。
- 要素和信息关联的表达集,最好列出每个命名关联的至少一个实例。(不必为可能由关联链接的每对对象类创建实例)
- 每个元要素和数据质量要素至少一个实例。

还应至少准备一个更新数据集,以验证更新数据集的格式和打包。

附加的测试数据集应该测试典型的数据量、代表性的数据获取问题和错误案例。

如果要以交换集的形式交付,则应将测试数据集打包为完整的交换集,包括示例元数据文件。对于其他形式的交付(事务、消息、Web服务等)的示例包,应该尽可能真实地进行仿真,如为Web服务设置、Web服务器、服务代理,但这些不是基本要求(它可以,并且应该作为测试阶段一部分)。

B-22 测试和反馈

应该准备一个正式的测试计划,包括测试用例。

B-23 工作流程

B-23.1 注册和请求 S 编号

申请产品规范编号:通常在开发项目启动阶段获得批准时指定规范的编号。
应按照 IHO 出版物 S-99 中规定的程序在 IHO GI 注册系统中注册产品规范。

B-23.2 项目组

项目组应包括领域专家、信息建模专家和原始设备制造商/开发方的代表。

B-23.3 作为开发过程的迭代优化

开发应制定迭代优化计划,并在指定的开发阶段审查以下内容:

- 初始化应用模式。对后续修订的审查可转到对主要产品规范文件的审查中;

- 产品规范和 DCEG 主要文件初稿。随后的修订应在准备就绪时进行审查；
- 每次对应用模式和要素目录进行重大修改后,都应检查要素目录、数据格式和样本数据集；
- 图示表达目录:初稿和重大修订；
- 其他制件或组件(例如有效性测试),在基本完成时,以及由于应用模式、要素目录或数据格式的更改而进行修改后；
- 将产品规范作为一个整体:在完整的包准备好之后。

对不同组成部分的审查可以结合起来,以符合开发时间表或工作组会议时间表。应用模式、主产品规范文档和 DCEG 应该在开发的不同阶段进行多次评审。

开发阶段的审查应向以下机构提出：

- 项目组:初步审查；
- 在达到一定的稳定性之后,由技术小组以及相关的技术工作组发起规范的编制工作；
- 开发人员、实施者和原始设备制造商:在项目团队和赞助技术工作组中进行初步审查和稳定后进行正式审查。请注意,单个实现者/原始设备制造商/开发人员最好从开发的最初阶段就参与进来。

评审应该是对完整性、正确性、获取和表达域的能力、性能/效率以及验证 S-100 和基础标准符合性所进行的一致性评审。

此外,IHO 数据质量工作组应该对产品规范进行完整性审查。当规范成熟时,应要求利益相关方进行审查,应包括：

- 生产者；
- 开发人员和原始设备制造商。

后续阶段应包括用户和用户测试,并编写一份影响研究报告。

测试开发和测试应从要素目录和数据格式开始,以达到合理的稳定性,可能需要经过一两个周期的审查。

IHO 的最终评估将在 HSSC 级别或其他同级别组织进行。

出版前审查将由 IHO 或其他出版机构在出版前进行,以检查出版问题。

B-23.4 产品规范维护

澄清、更正和修订应按照 S-100 第 12 部分第 12-2 条(维护程序)中所述的与 S-100 使用的相同标准指定。

产品规范应定期审查,建议新规范的审查周期为两年,可在规范成熟后延长至五年。

C 部分 数据质量

C-1 概述

S-100 通用海道测量数据模型是一种海道测量地理空间数据标准,可以支持各种海道测量相关的数字数据源,并与主流国际地理空间标准,特别是 ISO 19000 系列地理标准完全一致。这种方式使海道测量数据和应用程序更容易集成到地理空间解决方案中。S-100 本质上比 S-57 更灵活,它可以使用图像和网格数据类型、增强的元数据和多种编码格式。它还通过专门的在线注册表为要素、属性和图示提供了更灵活和动态的维护机制。S-100 提供了一个组件框架,能够为海道测量数据建模建立标准化的产品规范,从而在不同的数据标准和系统之间提供真正的互操作性。

S-97 是一本指南,目标读者是使用 IHO S-100 标准(通用海道测量数据模型)产品规范的开发人员和维护人员。

C-2 介绍

C 部分旨在确保所有基于 S-100 的产品规范都以适当和协调的方式处理数据质量方面的问题。

本数据质量指南可供开发基于 S-100 的产品规范的 HSSC 工作组使用。它为基于 S-100 的产品规范的数据质量度量的适用性提供了 10 条建议。

起草产品规范时,数据质量指南将作为指导文件,以验证产品规范中是否包含了适当的数据质量元素。数据质量元素是记录数据集质量的定量组件。数据质量元素对数据集的适用性取决于数据集的内容及其产品规范,数据质量元素不可能适用于所有数据集。

各种 ISO 标准(ISO 19115、ISO 19139、ISO 19115-1/2/3、ISO 19157)和 S-100 第 4a ~ 4c 部分描述了数据集和交换集元数据中数据质量度量的位置以及元数据中数据质量的编码。

C-3 规范性引用文件

IG-D2.8. II.1 D2.8. II.1 INSPIRE 高程数据规范 技术指南。

ISO 8211 信息交换结构实施用数据描述文件规范。ISO/IEC 8211,1994 年。

ISO 19115 地理信息 元数据(2003),经2006年第1号修正案修正。

ISO 19115-1 地理信息 元数据 第1部分:基础,ISO19115-1,2014年,经2018年第1号修正案修订。

ISO 19115-2 地理信息 元数据 第2部分:图像和网格数据的扩展,ISO 19115-2,2009年。

ISO 19115-3 地理信息 元数据 基本概念的XML模式实现,ISO/TS 19115-3,2016年。

ISO 19139 地理信息 元数据 XML模式实现。

ISO 19157 地理信息 数据质量,ISO 19157,2013年,经2018年第1号修正案修正。

S-44 IHO S-44 海道测量规范

S-100 IHO S-100 通用海道测量数据模型,4.0.0版,2018年12月。

S-101 IHO S-101 电子海图产品规范,1.0.0版,2018年12月。

S-102 IHO S-102 水深表面产品规范,1.0.0版,2012年4月。

S-121 IHO S-121 海上边界与界限,1.0.0版,2019年10月。

注:除非特别标明版本,本文档提及的S-100均指S-100 4.0.0版

C-4 术语和缩略词

C-4.1 术语

数据产品规范 data product specification

数据集或数据集系列的详细描述,以及使其能够被另一方创建、提供和使用的附加信息。数据产品规范提供了对海道测量概念的描述,并提供了将讨论范围映射到数据集的规范。可用于生产、销售、终端使用或其他目的。

数据质量元素 data quality element

记录数据集质量的定量成分。[ISO 19101:2002;S-100 4.0.0 附件A]。

注:数据质量元素对数据集的适用性取决于数据集的内容及其产品规范,但数据质量元素不可能适用于所有数据集。

数据质量评估程序 data quality evaluation procedure

应用和报告质量评价方法和结果的全部操作。

数据质量度量 data quality measure

数据质量子元素的评估。

数据质量概述元素 data quality overview element

记录数据集质量的非定量组件。有关数据集的用途,使用和沿袭的信息是非定量的质量信息。

数据质量结果 data quality result

通过应用数据质量度量而产生的值或值集,或根据指定的一致性质量级别对所获得的值或值集进行评估的结果。

数据质量范围 data quality scope

用于报告质量信息的数据的范围或特征(多个)。

注:数据集的范围可以包括数据集所属的数据集系列、数据集本身或者物理上位于数据集中的共享公共特性的较小数据组。公共特征可以是被识别的要素类型、要素属性或要素关系、数据收集标准、原始资料来源或特定的地理或时间范围。

数据质量子元素 data quality sub-element

描述数据质量元素某一方面的数据质量元素的组件。

元数据 ISO 19115 Metadata ISO 19115

符合 ISO 相关标准(ISO 19115-1/2/3 和 ISO 19157)规定格式的质量报告。

元质量 metaquality

有关数据质量结果可靠性的信息。[ISO 19157]

结果范围 result scope

(数据质量)结果的范围。[ISO 19157, 改编]

注:结果范围是数据质量范围的子集。

独立质量报告 standalone quality report

提供关于所用数据质量评估、结果和措施的全面的详细信息的自由文本文件。[ISO 19157:2013]

C-4.2 缩略词

DQWG 数据质量工作组(IHO Data Quality Working Group)

ECDIS 电子海图显示与信息系统(Electronic Chart Display and Information System)

ECS 电子海图系统(Electronic Chart System)

ENC 电子海图(Electronic Navigational Chart)

GI 地理空间信息(Geospatial Information)

GML 地理标记语言(Geography Markup Language)

GSD 地面采样距离(Ground Sampling Distance)

HDF5 分层数据格式第5版(Hierarchical Data Format Version 5)

IEC 国际电工协会(International Electrotechnical Commission)

IHO 国际海道测量组织(International Hydrographic Organization)

IMO 国际海事组织(International Maritime Organization)

ISO 国际标准化组织(International Organization for Standardization)

PS 产品规范(Product Specification)

RMSE 均方根误差(Root Mean Square Error)

SD 标准差(Standard Deviation)

XML 可扩展标记语言(eXtensible Markup Language)

C-5 数据质量度量概述

数据质量项之间的关系如图 C-5-1 所示。术语已在 C-4.1 条中定义。

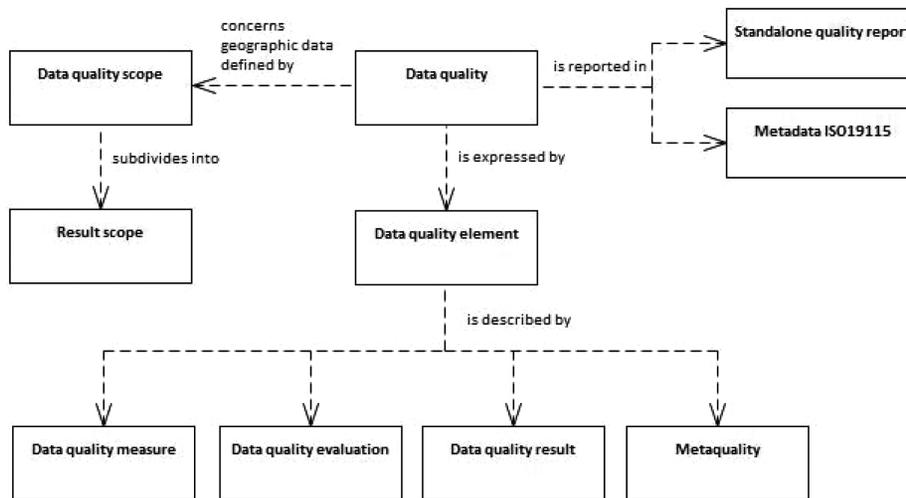


图 C-5-1 地理数据质量概念模型⁶

数据质量度量的组件可分为以下元素⁷：

1. 完整性
2. 逻辑一致性
3. 位置精度
4. 专题准确性
5. 时间质量
6. 聚合
7. 可用性

数据质量评估可分为以下要素：

1. 数据质量全面检查
2. 数据质量抽样检验
3. 数据质量间接评价
4. 数据质量聚合派生

数据质量结果可分为以下元素：

1. 数据质量一致性结果

⁶ ISO 19157 地理信息 数据质量,第6页。

⁷ ISO 19157 地理信息 数据质量,第7页。

2. 数据质量定量结果
3. 数据质量描述结果

C-5.1 数据质量度量

本节对数据质量度量的描述引自 ISO 19157。

完整性被定义为要素、要素属性和关系的存在和不存在。它由两个数据质量元素组成：

- 冗余：数据集中存在的多余数据；
- 缺失：数据集中缺少的数据。

逻辑一致性被定义为对数据结构、属性和关系(数据结构可以是概念的、逻辑的或物理的)的逻辑规则的遵守程度。如果这些逻辑规则在其他地方有文档记录(例如在产品规范中),那么应该引用源(例如在数据质量评估中)。它由四个数据质量元素组成：

- 概念一致性：对概念模式规则的遵守；
- 域一致性：值与值域的一致性；
- 格式一致性：根据数据集的物理结构存储数据的程度；
- 拓扑一致性：数据集显式编码拓扑特性的正确性。

位置精度被定义为要素在空间参考系统中的位置精度。它由三个数据质量元素组成：

- 绝对或外部精度：报告的坐标值与被接受为真或为真值的接近程度；
- 相对或内部精度：要素在数据集中的相对位置与其各自被接受为真或为真相对位置的接近程度；
- 网格数据位置精度：网格数据空间位置值与被接受为真或为真值的接近程度。

主题准确度是指定量属性的准确度、非定量属性的准确度、要素分类及其关系的准确度。它由三个数据质量元素组成：

- 分类正确性：将分配给要素或其属性的类别与语篇范围(例如基本真理或参考数据)进行比较；
- 非定量属性的正确性：衡量一个非定量属性是否正确的尺度；
- 定量属性准确度：定量属性值与被接受为真或已知为真值的贴适度。

时间质量是指要素的时间属性和时间关系的质量。它由三个数据质量元素组成：

- 时间测量的精确性：报告的时间测量值与被接受为真或已知为真值的接近程度；

- 时间一致性:事件顺序的正确性;
- 时间有效性:数据相对于时间的有效性。

可用性基于用户需求。所有质量元素都可以用来评估可用性。可用性评估可以基于不能使用上述质量元素描述的特定用户需求。在这种情况下,可用性元素应用于描述关于数据集是否适合特定应用或是否符合一组要求的特定质量信息。

C-6 给产品规范开发人员的建议

S-97 C 部分将 ISO 19157 中的数据质量概念应用于基于 S-100 的产品规范的开发。本文件提供了十项建议,有助于找到和应用 S-100 中要求的适用数据质量度量。

C-6.1 完整性 > 冗余/缺失

DQ_CompletenessCommission:S-100 附录 4c-C 中定义的海道测量质量元数据属性定义。

DQ_CompletenessOmission:S-100 附录 4c-C 中定义的海道测量质量元数据属性定义。

建议 1:数据质量度量完整性(冗余/缺失)应包括在产品规范中。

C-6.2 逻辑一致性 > 概念一致性

S-100 第 1 部分中对概念模式语言给出了定义,提供了如下描述:

- 类;
- 属性;
- 基本数据类型;
- 几何图元类型;
- 复杂类型;
- 预定义派生类型;
- 枚举类型;
- 代码表类型;
- 关系和关联;
- 组合和聚合;
- 原型以及可选、有条件和强制的属性和关联;
- 命名和命名空间;
- 注释;和

- 包。

建议 2:数据质量量度概念一致性应遵循 S-100 第 1 部分中的准则,并纳入产品规范。

C-6.3 逻辑一致性 > 域一致性

这一部分在 S-100 第 5 部分要素目录中进行了描述。该部分提供了使用地理数据集组织和报告现实世界地理现象的标准框架。它定义了要素类型和分类方法,以及在要素目录中组织并呈现给用户的方法。该方法适用于在以前未编目的领域中创建要素类型的目录,并适用于修改现有的要素目录以符合标准实践。它适用于以数字形式表示的要素类型的编目。其原则可以扩展到其他形式的地理数据的编目。

建议 3:数据质量量度域一致性应遵循 S-100 第 5 部分的指南,并纳入产品规范。

C-6.4 逻辑一致性 > 格式一致性

这部分在 S-100 第 10 部分编码格式中进行了描述。S-100 未强制规定编码格式,因此产品规范开发者可以决定适合的编码标准并指定具体编码格式。这主要是由于编码类型非常多,包括且不局限于 ISO/IEC8211、GML、XML、GeoTiff、HDF-5、JPEG2000 等。

建议 4:数据质量度量格式一致性应遵循 S-100 第 10 部分的指南,并纳入产品规范。

C-6.5 逻辑一致性 > 拓扑一致性

这部分在 S-100 第 7 部分空间模式中进行了描述。它支持 0、1、2 和 2.5 维的空间模式和两种复杂度级别——几何原型和几何复型。

拓扑一致性的条件在 S-100 第 7 部分空间模式 7-4.3 节和附录 7-A 中进行了描述。详情参见 7-4.3 和附录 7-A 中的图。

建议 5:数据质量衡量拓扑一致性应遵循 S-100 第 7 部分中的指南,并纳入产品规范。

C-6.6 位置精度

位置精度由 S-100 第 4c 部分 元数据 数据质量描述。

它可以进一步细分为绝对或外部精度、垂直位置精度、水平位置精度和网格数据位置精度。

应当注意应用空间数据参考的不同方式。点集数据包括点集中每个点(点/曲线)的坐标直接位置。网格化数据将网格作为一个整体引用。网格化数据的两个空间属性描述了空间范围如何被划分为小单元和对地球的空间参考。ISO 19123 标准指示网格可以根据坐标参考系统来定义。这需要关于网格原点在坐标参考系统内的位置、网格轴的方向以及网格线之间间距的度量的附加信息。以这种方式定义的网格称为纠正网格。如果坐标参考系统是通过一个基准面与地球相关的,那么这个网格就是地理校正网格。其实质是网格坐标到外参照系坐标的一个仿射变换。

目前在海道测量领域,95%的置信度(高斯分布)是常用的位置精度。均方根误差(RMSE)是科学界常用的一种误差估计方法。RMSE是数据集坐标值和坐标值之间的平方差集合的平均值的平方根,坐标值来自对同一点具有更高精度的独立源。

根据具体的产品规范,也可以采用其他计算方法。对 S-44、S-101、S-102 和 S-121 进行了比较。它们使用不同的计算方法和/或表达相同的概念(不确定性)。使用的方法如下:

- S-44:95% ($2 \times SD$)
- S-121:标准循环错误 $\{ = 0.7071 \times \text{SQRT}[SD(X) + SD(Y)] \}$,已转换为类别属性。
- S-101:所有水平位置(2D)、垂直(1D)、水平距离(1D)和方向(1D)不确定性属性都涉及所有测量、处理和可视化误差源相关的变化的95%置信度。不确定度编码为:(1)不确定度固定[一维误差(垂直方向)或二维误差(水平方向)的最大绝对值,假定误差为正负];(2)不确定变量因子[应用于计算某一量的不确定度的因数。表示与导致可变不确定度分量的因数(或百分比)相等的分数,即5%的因数编码为0.05。]

上述情况就像是在不同的产品规范之间分别使用海里、英里和公里。这在单独使用时可能很好,但是一旦开始组合不同的产品规范并使用计算机算法来创建一个基于这些质量参数的“智能”系统,则开发系统时必须非常小心,以确保不会混淆。

IHO DQWG 正在考虑解决这种情况的不同方法,可能会让单独的产品规范保留自己的参数,但需要告知开发人员和规范的其他用户如何从一种精度标准转换到另一种精度标准。

建议 6:数据质量度量位置精确度应遵循 S-100 第 4C 部分的指南,并纳入产品规范。位置精度的计算应尽可能与其他基于 S-100 的产品规范协调一致。

C-6.7 主题准确性

在 ISO 19157 中进行了描述。

进一步划分为三个部分:定量属性精度、非定量属性精度和专题分类准确性。

地理应用程序的数据内容是根据真实世界要素的图和特定应用程序在文中的要求定义的。内容按照对象进行结构化。本文考虑两种类型的对象:

- 1) 要素:要素与其属性一起定义。
- 2) 信息类型:信息类型用于在要素和其他信息类型之间共享信息。信息类型仅具有主题属性。对于主题分类的准确性,将一个项分配给某个类可能是正确的,也可能是不正确的。对于定量属性的准确性,可以用不确定度区间来衡量定量属性的准确性。对于非定量属性的准确性,非定量属性的属性值可以是正确的,也可以是不正确的。

建议 7:数据质量度量专题准确性应遵循 S-100 第 4c 部分的准则,并纳入产品规范。

C-6.8 时间质量

时间质量由第 4c 部分 元数据 数据质量进行了描述。

建议包括时间一致性和时间有效性,因为这为用户提供了在时间与空间域中正确注册信息的保证。对于具有非常精确的时间属性(例如遥感)的数据元素,也可以提供时间精度。

建议 8:数据质量度量时间质量应遵循 S-100 第 4c 部分的准则,并纳入产品规范。应包括时间一致性和时间有效性。

C-6.9 聚合

上层指定的数据质量(例如系列)适用于下层(例如数据集),见下表 C-6-1(引用 ISO 19157)。如果上层和下层之间的数据质量不同,则应在下层提供补充信息。

层次结构

表 C-6-1*

上层 ↓ 下层	系列	
	数据集	
	子集	
	要素类型	属性类型
	要素实例	属性实例

在聚合不同质量结果的情况下,独立质量报告将提供关于原始结果的完整信息(采用评价程序和措施)、聚合结果和聚合方法,而元数据仅描述聚合结果,可以参考独立质量报告中描述的原始结果。

如果数据集已传递了对数据产品规范的一致性,则聚合数据质量结果将提供一个结果。

建议 9:数据质量度量聚合结果应被包含在指明数据集/数据及系列是否传递给产品规范。

C-6.10 可用性

可用性基于用户需求。所有质量元素应用于评估可用性。可用性评估可以基于那些无法用上述质量元素描述的特定需求。在这种情况下,可用性元素可以用来描述数据集对于特殊应用和需求的特定质量信息。

所有产品规范都应有一个描述数据质量的图。为了确保不同产品规范间的协调,DQWG 推荐所有产品规范都共享一个通用文本来解释数据质量的概念,见下文的数据质量介绍。下文是关于这一共同介绍的提案:

数据质量介绍

数据质量允许用户和用户系统评估所提供的数据的使用情况。数据质量度量和相关评估被报告为数据产品的元数据。此元数据提高了与其他数据产品的互操作性,并提供了不适用数据产品的用户群。次要用户可以根据报告的数据质量度量评估数据产品在其应用中的有用性。

对于<本产品规范>,包含下列数据质量元素⁸:

- 遵循本产品规范;
- 数据产品的预期用途;

⁸ IHO 海道测量标准与服务委员会认为必要的。

* 译者注:此外疑原文有误,应为表 C-6-1*。

- 数据产品在覆盖方面的完整性;
- 逻辑一致性;
- 位置不确定度和精度;
- 专题准确性;
- 时间质量;
- 聚合度量;
- 验证检查或一致性检查,包括:
 - ✓ 数据集完整性的一般测试;
 - ✓ 特定数据模型的特定测试。

建议 10:所有基于 S-100 的产品规范将“数据质量介绍”段落用作模板。

C-7 数据质量度量

此列表源自 ISO 19157。

这里定义了 20 种不同的数据质量度量,可用于验证基于 S-100 的产品规范。如表 C-7-1 所示。下表建议了等深线和网格水深测量的位置精度目标结果。

适用于基于 S-100 的产品规范的 ISO 19157 度量的完整清单见 S-100 附录 4c-C 海道测量质量元数据属性定义。

推荐的数据质量度量

表 C-7-1

数据质量度量	定义	数据质量测量/描述	评价范围	空间表达类型 适用性
完整性/冗余	根据范围,数据集中的过量数据	numberOfExcessItems/该数据质量度量是不应该出现在数据集中的项的数量	数据集/数据集系列	所有基于 S-100 的 PS
完整性/冗余	根据范围,数据集中的过量数据	numberOfDuplicateFeatureInstances/该数据质量度量是数据中要素实例重复的数量	数据集/数据集系列	所有基于 S-100 的 PS
完整性/缺失	根据范围,数据集中的数据缺失	numberOfMissingItems/该数据质量度量是一个指示符,显示了数据中缺失的特定项	数据集/数据集系列/空间对象类型	所有基于 S-100 的 PS
逻辑一致性/概念一致性	与概念模式规则的符合性	numberOfInvalidSurfaceOverlaps/该数据质量度量是数据集中所有错误重叠的总数量。面可能重叠,但不可以是依赖应用的。不是所有的重叠面都是错误	空间对象/空间对象类型	含有几何面的 PS

续上表

数据质量度量	定义	数据质量测量/描述	评价范围	空间表达类型 适用性
逻辑一致性/域一致性	值与值域的符合性	numberOfNonconformantItems/该数据质量度量是数据集中与值域不符的所有项	空间对象/ 空间对象 类型	所有基于 S-100 的 PS
逻辑一致性/格式一致性	根据范围,数据存储与数据集物理结构的符合性	physicalStructureConflictsNumber/该数据质量度量是数据集中与数据集物理结构冲突的项的总和	数据集/数 据集系列	所有基于 S-100 的 PS
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	rateOfFaultyPointCurveConnections/该数据质量度量是链接节点的错误数,提供了错误的点-曲线连接与总的点-曲面连接数量的关系	空间对象/ 空间对象 类型	含有曲线的 PS
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	numberOfMissingConnectionsUndershoots/该数据质量度量是数据集的容差中未达到参数的项的数量	空间对象/ 空间对象 类型	含有曲线的 PS
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	numberOfMissingConnectionsOvershoots/该数据质量度量是数据集的容差中超过参数的项的数量	空间对象/ 空间对象 类型	含有曲线的 PS
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	numberOfInvalidSlivers/该数据质量度量是数据集中无效的碎面项的数量。碎片是由于相邻面未进行有效数字化造成的。相邻面的边界可能存在空隙或重叠,从而引起拓扑错误	数据集/数 据集系列	含有几何面 的 PS

续上表

数据质量度量	定义	数据质量测量/描述	评价范围	空间表达类型 适用性
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	numberOfInvalidSelfIntersects/该数据质量度量是数据集中非法的交叉项的总和	空间对象/ 空间对象 类型	含有曲线/几何 面面的 PS
逻辑一致性/拓扑一致性	根据范围,数据集显示编码的拓扑特征的正确性	numberOfInvalidSelfOverlap/该数据质量度量是数据集中非法的自重叠项的总和	空间对象/ 空间对象 类型	含有曲线/几何 面面的 PS
位置精度/绝对 或外部精度	报告坐标值与真值之间的 接近程度	均方根误差/标准差,真值并非从观测得来, 而是先前已知的	空间对象/ 空间对象 类型	含有物标坐标 位置信息值的 PS
位置精度/垂直 位置精度	报告坐标值与真值之间的 接近程度	linearMapAccuracy2Sigma/上下限定义的间隔 的一半长度,其中真实值的概率为 95%	空间对象/ 空间对象 类型	含有垂直坐标 位置信息值的 PS
位置精度/水平 位置精度	报告坐标值与真值之间的 接近程度	linearMapAccuracy2Sigma/上下限定义的间隔 的一半长度,其中真实值的概率为 95%	空间对象/ 空间对象 类型	含有水平坐标 位置信息值的 PS
位置精度/网格 数据位置精度	报告坐标值与真值之间的 接近程度	平面测量的均方根误差/特定圆点构成的圆 的半径,真实值的概率为 P	空间对象/ 空间对象 类型	含有网格坐标 位置信息值的 PS
时间质量/时间 一致性	时间一致性	有序事件或序列的正确性	数据集/数 据集系列/空 间对象类型	含有时间价值 的 PS

续上表

数据质量度量	定义	数据质量测量/描述	评价范围	空间表达类型 适用性
主题精度/主题 分类正确性	分配给要素的类或它们的 属性与话语世界的比较	miscalculationRate/该数据质量度量是不正确 的要素类的数量。(引自 ISO 19157)这是一个 比率,使用实数分数形式表示。例如,如果有 1 项被错误的分类,数据集中有 100 个项,那么比 率为 1/100,即 0.01	数据集/数 据集系列/空 间对象类型	所有基于 S-100 的 PS
聚合度量/聚合 度量	在数据产品规范中,有一些 特定的数据产品需求	DataProductSpecificationPassed/该数据质量度 量是一个布尔值,指示产品规范中所有数据的 需求都已满足	数据集/数 据集系列/空 间对象类型	需要完全传递 数据集/数据集系 列/空间对象数据 类型所有元素 的 PS
聚合度量/聚合 度量	在数据产品规范中,有一些 特定的数据产品需求	DataProductSpecificationFailRate/该数据质量 度量是一个数字,表示当前产品/数据集不满足 数据产品规范数据需求的总体数量	数据集/数 据集系列/空 间对象类型	需要完全传递 数据集/数据集系 列/空间对象数据 类型所有元素 的 PS

注:位置精度/绝对或外部精度建议⁹:

$$\text{最大 RMSE(水平)} = E/10000$$

$$\text{最大 RMSE(垂直)} = V_{int}/6$$

位置精度/网格数据位置精度建议:

$$\text{最大 RMSE(水平)} = GSD/6$$

$$\text{最大 RMSE(垂直)} = GSD/3$$

E = 预期编绘比例尺分母

⁹ D2.8. II.1 高程数据规范 技术指南,第 95 页。

其中:

Vint = 标准等深线间隔

GSD = 地面采样距离

C-8 数据验证的最低标准

数据验证的最低标准检查集正在开发中,完成后将添加到此文档中。