

中华人民共和国国家标准

GB/T 25000.12—2017

系统与软件工程 系统与软件质量要求和 评价(SQuaRE) 第12部分：数据质量模型

**Systems and software engineering—Systems and software Quality Requirements
and Evaluation(SQuaRE)—Part 12:Data quality model**

[ISO/IEC 25012:2008,Software engineering—Software product Quality
Requirements and Evaluation(SQuaRE)—Data quality model,MOD]

2017-11-01 发布

2018-05-01 实施

中华人民共和国国家质量监督检验检疫总局 发布
中国国家标准化管理委员会

目 次

前言	Ⅲ
引言	V
1 范围	1
2 符合性	2
3 规范性引用文件	2
4 术语和定义	2
5 数据质量	3
5.1 概述	3
5.1.1 固有的数据质量	3
5.1.2 依赖系统的数据质量	4
5.2 数据质量模型	4
5.3 数据质量特性	5
5.3.1 固有的视点	5
5.3.2 固有的且依赖系统的视点	6
5.3.3 依赖系统的视点	8
附录 A (资料性附录) 来自GB/ T 25000.1—2010的术语和定义	10
参考文献	11

前 言

GB/T 25000《系统与软件工程 系统与软件质量要求和评价(SQuaRE)》分为如下几部分：

- 第1部分：SQuaRE 指南；
- 第2部分：计划与管理；
- 第10部分：系统与软件质量模型；
- 第12部分：数据质量模型；
- 第20部分：测量参考模型和指南；
- 第21部分：质量测度元素；
- 第22部分：使用质量测量；
- 第23部分：系统和软件产品质量测量；
- 第24部分：数据质量测量；
- 第30部分：质量需求；
- 第40部分：评价过程；
- 第41部分：开发方、需方和独立评价方的评价指南；
- 第42部分：评价模块；
- 第45部分：可恢复性的评价模块；
- 第51部分：就绪可用软件产品(RUSP) 的质量要求和测试细则；
- 第60部分：易用性测试报告行业通用格式(CIF)：易用性相关信息的通用框架；
- 第62部分：易用性测试报告行业通用格式(CIF)；
- 第63部分：易用性的行业通用格式(CIF)：使用周境描述；
- 第64部分：易用性的行业通用格式(CIF)：用户要求报告；
- 第65部分：易用性的行业通用格式(CIF)：用户需求规格说明；
- 第66部分：易用性的行业通用格式(CIF)：评价报告。

本部分是GB/T 25000 的第12部分。

本部分按照GB/T 1.1—2009 给出的规则起草。

本部分采用重新起草法修改采用 ISO/IEC 25012:2008《软件工程 软件产品质量要求和评价 (SQuaRE) 数据质量模型》，与ISO/IEC 25012:2008相比，主要技术差异如下：

- 关于规范性引用文件，本部分做了具有技术性差异的调整，以适应我国的技术条件，调整的情况集中反映在第2章“规范性引用文件”中，具体调整如下：

- 用等同采用国际标准的 GB/T 25000.1—2010 代替了ISO/IEC 25000。

- 为保持本部分术语与ISO/IEC 25024:2015《系统与软件工程 系统与软件质量要求和评价 (SQuaRE) 数据质量测量》的一致性，参照ISO/IEC25024:2015 对本部分术语4.1、4.2、4.7、4.10、4.14进行了修订。

本部分做了下列编辑性修改：

- 根据最新 SQuaRE 系列国际标准的规范，参照ISO/IEC 25024:2015 对引言中有关 SQuaRE 系列国际标准的介绍进行了修订。
- 根据最新 SQuaRE 系列国际标准的规范，对本部分标准名称进行了修改。

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这些专利的责任。

本部分由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本部分起草单位：万达信息股份有限公司、复旦大学、招商银行股份有限公司、东莞市雅康精密机械有限公司、佛山柯维光电股份有限公司、珠海力准电子科技有限公司、中国电子技术标准化研究院、中国移动(深圳)有限公司、国家应用软件产品质量监督检验中心、上海计算机软件技术开发中心、上海鸿迪信息技术有限公司、招银云创(深圳)信息技术有限公司、上海市软件评测中心有限公司、重庆市软件评测中心有限公司、四川省软件和信息系统工程测评中心、浙江省电子信息产品检验所、湖北软件评测中心、福建省电子产品监督检验所、吉林省电子信息产品监督检验研究院。

本部分主要起草人：李光亚、吴毅坚、钱乐秋、赵文耘、张旻旻、陈诚、肖筱华、王志鹏、刘潇健、蔡立志、高旭磊、徐鸿俊、贾俊刚、舒敏根、李莉、何志明、周清云、孟艳、何志峰、张凯、罗国翔、胡芸、王威、陈振宇、丁晓明、冯丽、季永炜、夏启明、柳毓龙、辛士界。

引 言

计算机系统处理的数据和信息量在全球范围内不断增长；数据质量是从这些数据导出信息的质量和有用性的关键要素，大多数业务过程都依赖于数据质量。

所有信息技术项目的一个共同前提条件是在计算机系统和用户之间以及计算机系统自身之间所交换、处理及使用的数据的质量。

管理并提高数据质量是重要的，这是由于：

- 从数据生产过程的质量未知或较差的组织获取数据；
- 存在有缺陷的数据，造成令人不满意的信息、无用结果和客户不满意；
- 这种数据在不同所有者和用户之间传播，按照单个组织的工作流需要获取的数据经常缺少对保证互操作性和合作所必需的一致和集成的视图；
- 需要处理由于语义歧义性或者与其他现有的相关数据间缺乏一致性而无法直接重复使用的数据；

——遗留体系结构和计算机系统的共存，这些计算机系统包含在不同时期、按照不同标准设计和实现的分布的系统；

——现存的信息系统(例如万维网)数据频繁变化，集成成为一个特别重要的问题。

本部分定义的数据质量模型旨在满足这些要求，并且考虑到数据生存周期通常比软件生存周期更长。该数据质量模型可以用于以下示例方面：

- 在数据生产、获取和集成过程中定义和评价数据质量需求；
- 标识数据质量保证准则，同时也有利于数据的再工程、评估和改进；
- 评价数据与法律和(或)需求的依从性。

检测因数据带来的错误或低效，带来了针对数据和数据所在系统的其他组件的改善和纠正性干预，示例如下：

- 数据(例如重新设计、解析、整理、增补、转换、匹配)；
- 软件(例如修改源程序以实现一致性控制)；
- 硬件(例如升级计算机系统以提高响应时间)；
- 人的业务过程(例如对用户培训以避免数据录入过程中的错误；改进管理数据的账目处理)。

本部分定义的数据质量模型，旨在与 SQuaRE 系列标准中的其他标准一起使用。

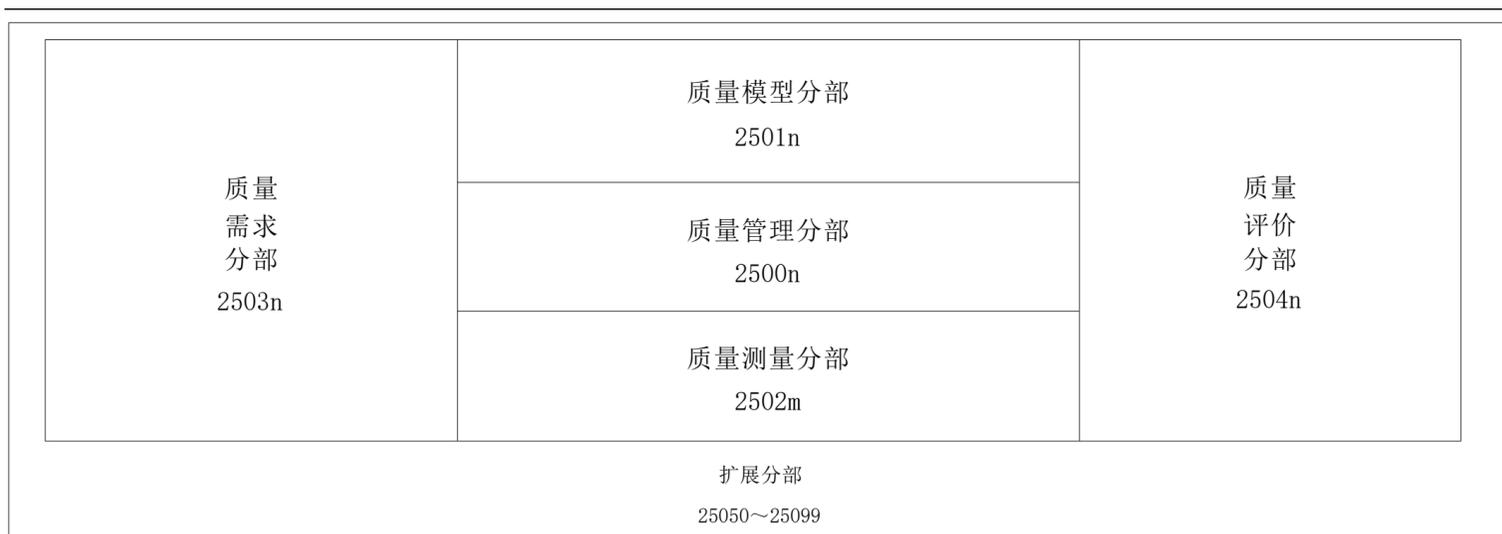


图 1 SQuaRE 系列国际标准的组织

SQuaRE 系列国际标准由五大主分部和扩展分部组成，如图1所示(改写自ISO/IEC 25000:2005)。SQuaRE 系列标准中各分部的概述如下。

- ISO/IEC 2500n——质量管理分部。构成这个分部的标准定义了由SQuaRE 系列标准中的所有其他标准引用的全部公共模型、术语和定义。这一分部还提供了用于计划和管理一个项目的需求和指南。
- ISO/IEC 2501n——质量模型分部。构成这个分部的标准为使用质量、数据以及系统或软件产品提供了的质量模型。服务质量正在开发中。还提供了使用该质量模型的实用指南。
- ISO/IEC 2502n——质量测量分部。构成这个分部的标准包括系统或软件产品质量测量参考模型、质量测量的定义及其应用的实用指南。这个分部给出了系统与软件产品质量测度、使用质量测度以及“固有的”视点和“依赖系统的”视点的测度。定义并给出了构成质量测量基础的质量测度元素。
- ISO/IEC 2503n——质量需求分部。构成这个分部的标准帮助用户规定质量要求。这些质量要求可用在要开发的系统或软件产品的质量需求抽取过程(设计一个达到必要质量的过程)中或用作评价过程的输入。
- ISO/IEC 2504n——质量评价分部。构成这个分部的标准给出了无论由独立评价方、需方还是由开发方执行的系统或软件产品评价的要求、建议和指南。还给出了作为评价模块的质量测度编制支持。

ISO/IEC 25050~ISO/IEC 25099 保留用于 SQuaRE 扩展国际标准，目前包括 ISO/IEC 25051 和 ISO/IEC 25060~ISO/IEC 25069。

系统与软件工程 系统与软件质量要求和评价(SQuaRE) 第12部分：数据质量模型

1 范围

GB/T 25000 的本部分针对计算机系统中以某种结构化形式保存的数据，定义了通用的数据质量模型。

本部分关注于作为计算机系统一个组成部分的数据的质量，并定义由人和系统使用的目标数据的质量特性。

目标数据是组织决定通过本模型来分析和确认的那些数据。非目标数据包括两种情况：第一种指非持久的数据，例如由操作系统处理的数据；第二种指虽然属于本部分范围、但该组织不采用本部分进行分析评价的数据。

图2中，显示了一个通用系统的结构：系统可以包含信息系统，信息系统又可以包含一个或多个计算机系统。

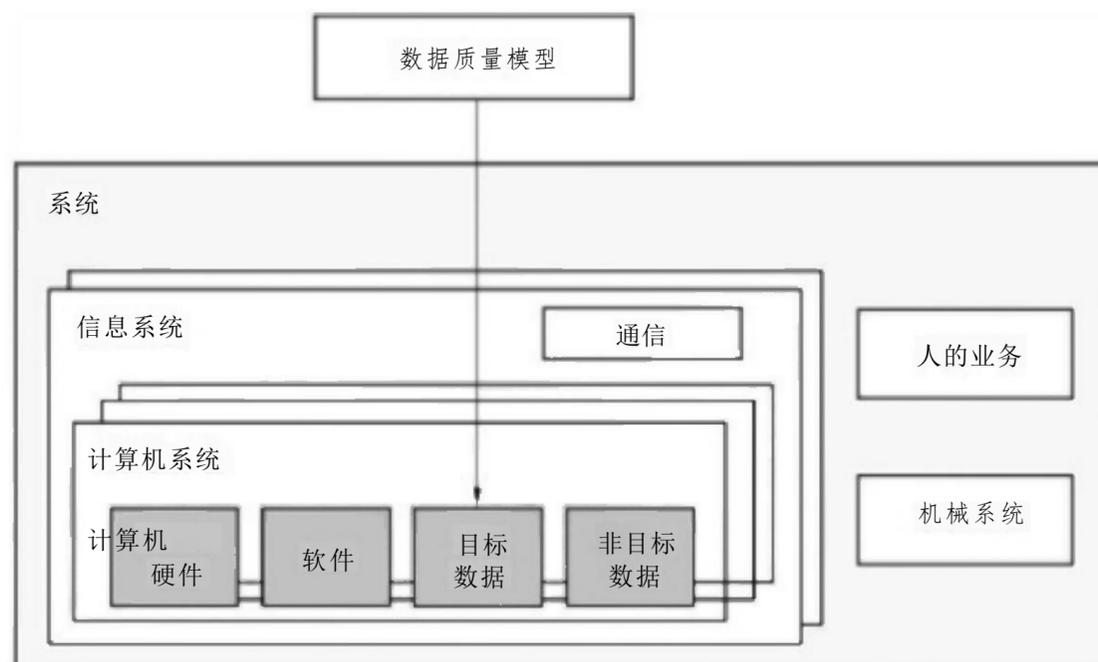


图 2 数据质量模型的目标域

本部分可以与SQuaRE 系列的其他标准共同用于建立数据质量需求、定义数据质量测度或规划并实施数据质量评价。

数据质量需求和数据质量测度能按5.2中的数据质量特性进行分类，并由评价过程以独立于其他计算机系统的组件的方式用于分析数据。

本部分旨在支持系统生存周期过程(例如 ISO/IEC 15288 中所定义的)的实现。

本部分考虑了所有数据类型(例如字符串、文本、日期、数字、图像、声音等)、所赋的数据值以及数据之间的关系(例如同一体或不同实体中数据之间的一致性)；范围不包括嵌入式装置或实时传感器所产生的、不保留作进一步处理或作为历史信息的数据。

本部分不对数据的物理组织(即数据库管理系统)作出规定；此外，概念、逻辑和物理模式设计活动也不在本部分的范围之内；与这些数据相关的所有处理和可交付物可以从本部分的应用中获益。

数据与数据设计之间的符合性包含在本部分的范围之内。

本部分为评价数据质量引用了元数据，但元数据的界定由GB/T 18391.1阐明。

本部分与工业和领域特定数据质量标准之间的关系以及对这些标准的优先级由用户在特定的使用周境中确定。

2 符合性

当声称符合本部分时，使用者应对5.2中的每个数据质量特性都提供已涉及的证据，或给出将某些特性排除在外的理由，或描述自己特有的数据质量属性分类并提供它们与5.2中特性的映射关系。

3 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是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 25000.1—2010 软件工程 软件产品质量要求和评价(SQuaRE) SQuaRE 指南(ISO/IEC 25000:2005, IDT)

4 术语和定义

GB/T 25000.1—2010界定的以及下列术语和定义适用于本文件。为了便于使用，本部分的附录 A 重复列出了GB/T 25000.1—2010 中的一些重点术语。

4.1

计算机系统 computer system

由一个或多个组件和元素[如计算机(硬件)、相关的软件和数据]组成的系统。

4.2

数据 data

信息的可再解释的形式化表示，以适用于通信、解释或处理。

注1:可以通过人工或自动手段处理数据。

[GB/T 5271.1—2000, 定义01.01.02]

注2:该定义没有完全引用GB/T25000.1—2010 的定义，因为后者是指与测量结果相关的数据。

4.3

数据质量 data quality

在指定条件下使用时，数据的特性满足明确的和隐含的要求的程度。

4.4

数据质量特性 data quality characteristic

对数据质量有影响的数据质量属性的类别。

4.5

数据质量测度 data quality measure

变量，为其赋值作为数据质量特性的测量结果。

注:改写ISO/IEC 15939:2007。

4.6

数据质量模型 data quality model

已定义的特性集合，提供一个框架用于说明数据质量需求和评价数据质量。

4.7

数据类型 data type

对属性的可能值集合、特性和操作集合的一种抽象的分类。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/797050130056006143>