

# 中华人民共和国国家标准

GB/T 41866.4—2022

## 系统与软件工程 信息技术项目绩效基准 度量框架 第4部分：数据收集和维护

Systems and software engineering—Information technology project  
performance benchmarking framework—Part 4: Data collection and  
maintenance

(ISO/IEC 29155-4:2016, Systems and software engineering—Information  
technology project performance benchmarking framework—Part 4: Guidance  
for data collection and maintenance, MOD)

(报批稿)

2022 - 10 - 14 发布

2023 - 05 - 01 实施

国家市场监督管理总局  
国家标准化管理委员会

发布

仅限内部学习使用

## 目 次

前 言 .....	III
引 言 .....	V
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 缩略语 .....	2
5 数据界定的要求和指南 .....	2
5.1 信息项目数据结构概述 .....	2
5.2 属性 .....	3
5.3 数据元素 .....	4
5.4 数据记录 .....	4
5.5 IT 项目数据集 .....	4
5.6 界定数据元素的要求 .....	5
5.7 界定数据记录的要求 .....	9
5.8 对界定 IT 项目数据集的要求 .....	9
6 收集数据功能 .....	10
6.1 概述 .....	10
6.2 “测量 IT 项目”中的数据收集功能 .....	10
6.3 “数据提交”中的数据收集功能 .....	10
6.4 “维护数据仓库”的数据收集功能 .....	11
7 数据收集的要求和指南 .....	11
7.1 数据源的要求和指南 .....	11
7.2 时限指南 .....	12
7.3 数据验证指南 .....	12
7.4 隐私保护的要求和指南 .....	12
8 维护与支持职能 .....	13
9 维护与支持的要求和指南 .....	13
9.1 数据维护的要求和指南 .....	13
9.2 数据仓库维护的要求和指南 .....	14
9.3 交付基准度量数据仓库产品的要求和指南 .....	14
9.4 交付已发布基准的要求与指南 .....	14
9.5 支持基准度量产品的要求和指南 .....	14
9.6 支持基准度量经验库的要求和指南 .....	15
附录 A （资料性） 基准度量信息模型 .....	16

附录 B （资料性） 数据元素分类和与其他有关标准的比较.....	23
附录 C （资料性） GB/T 41866.1-2022 中 IT 项目绩效基准度量框架模型 .....	29
参 考 文 献.....	31

仅限内部学习使用

## 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。

本文件是GB/T 41866《系统与软件工程 信息技术项目绩效基准度量框架》的第4部分。GB/T 41866已经发布了以下部分：

- 第1部分：概念和定义；
- 第2部分：基准度量要求；
- 第3部分：报告编制；
- 第4部分：数据收集和维护。

本文件修改采用ISO/IEC 29155-4:2016《系统与软件工程 信息技术项目绩效基准度量框架 第4部分：数据收集和维护指南》。

本文件与ISO/IEC 29155-4:2016的技术差异及其原因如下：

- 用规范性引用的GB/T 41866.1替换了ISO/IEC 29155-1（见第3章），以适应我国的技术条件；
- 用规范性引用的GB/T 41866.2替换了ISO/IEC 29155-2（见第3章），以适应我国的技术条件；
- 删除了ISO和IEC网址信息（见ISO/IEC 29155-4:2016的第3章），以适应我国的技术条件；
- 增加了“所提交数据的方式（例如，数据采用U盘、光盘等存储媒体或专用网络进行提交）”[见7.4a)]，以适应我国的技术条件；
- 增加了“并对存储媒体进行格式化处理”[见9.3a)]，以适应我国的技术条件，更全面地保护隐私信息。

本文件做了下列编辑性改动：

- 将标准名称调整为《系统与软件工程 信息技术项目绩效基准度量框架 第4部分：数据收集和维护》，以便表述更加准确；
- 用资料性引用的GB/T 8566—2022替换了ISO/IEC 12207:2008（见5.6.13中表3、表4），以适应我国的情况；
- 用资料性引用的GB/T 20917—2007替换了ISO/IEC 15939:2007（见附录A、附录B）；
- 用资料性引用的GB/T 22032替换了ISO/IEC 15288（见5.6.13中的注）；
- 资料性引用的GB/T 25000.21—2019代替了ISO/IEC 25021:2012（见附录B）；
- 增加了“定类尺度”“定序尺度”“定距尺度”“定比尺度”等数据元素的信息（见B.2.1.3），便于增加基准度量数据收集和维护的可操作性。

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。

本文件由全国信息技术标准化技术委员会（SAC/TC 28）提出并归口。

本文件主要起草单位：中国电子技术标准化研究院、深圳赛西信息技术有限公司、中国电子技术标准化研究院华东分院、山东山科数字经济研究院有限公司、山东省计算中心（国家超级计算济南中心）、北京软件和信息服务业交易所有限公司、北京软件造价评估技术创新联盟、成都淞幸科技有限责任公司、中国地质科学院地质力学研究所、北京中基数联科技有限公司、重庆市软件评测中心有限公司、中国航天系统科学与工程研究院、韶关市智慧科技有限公司、厦门理工学院、无锡市信息化协会、上海计算机软件技术开发中心。

本文件主要起草人：张旻旻、于秀明、张骏极、李文鹏、李敏、周鸣乐、吴春雷、郑旭飞、李刚、庄园、陆璐、李勇、张星星、于铁强、丁晓明、翟梦然、陈聪、吕雪、景清、曹新彬、韩德隆、卢俊文、陈广清、李世彬、王翔、王海青、胡芸、许宗敏、孙荣荣。

仅限内部学习使用

## 引 言

基准度量是指各“关注对象”相互比较或分别与一个基准进行比较，以此来评价该对象特性的活动。在GB/T 41866中，“关注对象”是指IT项目的绩效，而特性指IT项目的某一方面，例如，生产率。

基准度量是IT项目管理领域中发展最快的技术之一。IT项目绩效基准度量实例的产生和发展出于多种原因，其中最常见有：

- a) 比较类似行业项目生产率的需要；
- b) 比较不同项目类型和技术生产率的需要；
- c) 改进IT开发过程时找到最有效目标的需要；
- d) 比较不同供应商间的生产率的需要；
- e) 改进项目管理成熟度的需要；
- f) 改进项目估计能力的需要。

开展IT项目绩效基准度量的有关试验文献表明实施基准度量失败率较高，同时统计数据也证明了这一点。基准度量失败主因在于测量目标的选定和业务目标的选定不一致，以及对项目级与项目组合管理级之间关系的理解有误。当测量和结果不一致时，团队仍需花费不必要的精力来收集项目数据，这将降低继续进行基准度量并使之制度化的动力。

如图1所示，GB/T 41866系列标准拟包含以下多个部分：

- 第1部分：概念和定义。提供了信息技术项目绩效基准度量的总体框架模型。它包含了成功识别、定义、选择、应用和改进基准度量所需的必要活动和组件，也提供了信息技术项目绩效基准度量的术语和定义；
- 第2部分：基准度量要求。规范了进行和/或支持一个组织成功完成基准度量活动所要求的任务；
- 第3部分：报告编制。规范了报告编制过程和典型报告内容的通用要求和指南；
- 第4部分：数据收集和维护。提供了为输入和维护基准度量数据仓库而收集信息技术项目数据的活动的通用要求和指南。

.....

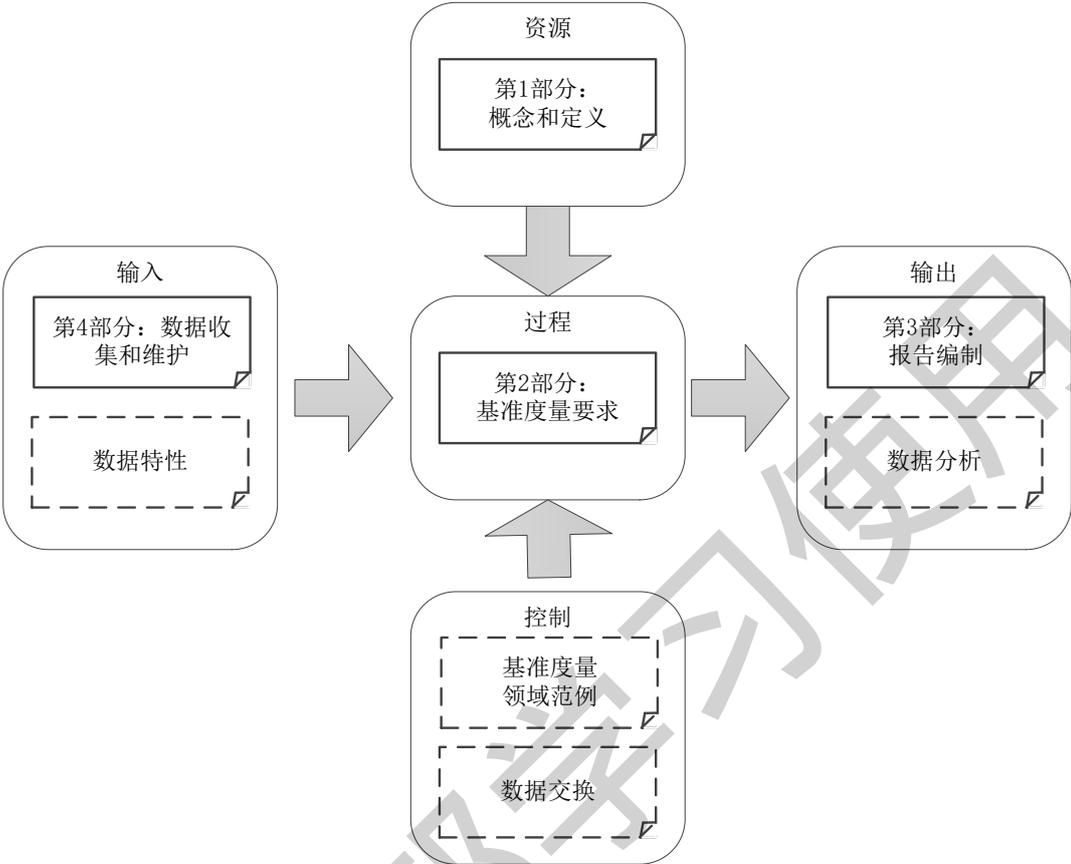


图1 IT项目绩效基准度量框架概览

# 系统与软件工程 信息技术项目绩效基准度量框架

## 第4部分：数据收集和维护

### 1 范围

本文件提供的通用的要求和指南，用于收集和维护信息技术（IT）项目数据，以及交付“IT 项目绩效基准度量框架”之内基准度量活动的基准数据：

- a) 数据元素界定的要求和指南；
- b) 基准度量框架内数据收集与维护过程的要求和指南；
- c) 维护基准度量数据仓库产品和已发布基准的要求和指南。

本文件主要关注三项重点活动，即“维护数据仓库”“提交 IT 项目数据”和“测量 IT 项目”。

本文件旨在供 IT 项目绩效基准度量的利益相关方（例如，基准度量用户、基准供方、基准度量服务供方和 IT 项目团队）使用。

注：本文件能如何使用的示例如下：

- 由基准供方界定数据元素，收集并维护 IT 项目数据，并提供基准度量数据仓库产品或已发布基准；
- 由基准度量分析方使用基准度量数据仓库的产品和/或基准，以执行基准度量的一个实例；
- 由基准度量服务供方利用基准度量数据仓库产品和/或基准，来提供基准度量服务；
- 由 IT 服务供方界定拟测量和/或拟提交给数据仓库拥有方的数据元素。

对于基准度量数据仓库中特别数据元素集的界定、格式或内容，均超出本文件范围。

### 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 41866.1 系统与软件工程 信息技术项目绩效基准度量框架 第1部分：概念和定义（GB/T 41866.1—2022，ISO/IEC 29155-1:2017, IDT）

GB/T 41866.2 系统与软件工程 信息技术项目绩效基准度量框架 第2部分：基准度量要求（GB/T 41866.2—2022，ISO/IEC 29155-2:2013, MOD）

### 3 术语和定义

GB/T 41866.1、GB/T 41866.2界定的以及下列术语和定义适用于本文件。

#### 3.1

**属性 attribute**

可由人或自动化工具定量或定性辨别的实体特征或特性。

[来源：GB/T 20917—2007, 3.2]

#### 3.2

**数据元素 data element**

IT 项目中最小的数据单位。

注1：数据元素通常作为基准度量数据仓库和/或“IT 项目数据”中一个数据单元来实现。

注2：数据元素由负责规划数据收集或基准度量的个人或组织来界定。

3.3

**数据记录 data record**

经明确界定而由有关数据元素（3.2）构成的组。其中将表示关注属性（3.1）的所有必要数据元素都包括进来。

注1：一个数据记录通常对应 IT 项目绩效基准度量框架内一个特定的 IT 项目。

3.4

**IT 项目数据集 IT project dataset**

依据预先界定的准则选出、收集起来，并带有类别的数据记录组。

注1：IT 项目数据集的分类准则可基于数据拥有方和/或用户的信息要求。

4 缩略语

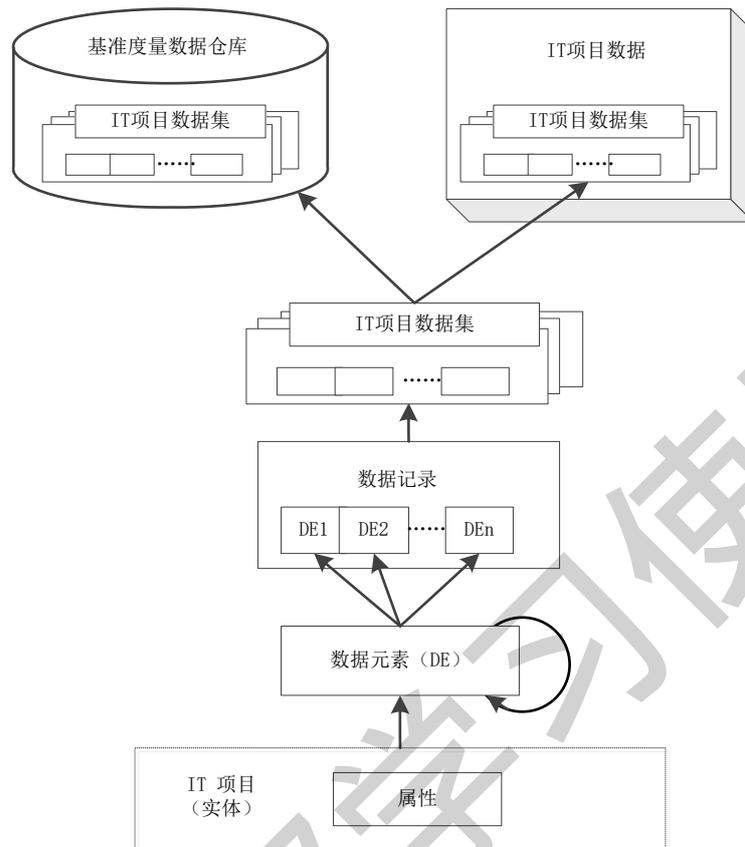
下列缩略语适用于本文件。

IT：信息技术（Information Technology）

5 数据界定的要求和指南

5.1 信息项目数据结构概述

GB/T 41866 系列的 IT 项目绩效基准度量框架中有两种主要的数据存储器，即“基准度量数据仓库”和“IT 项目数据（库）”（见图 C.1）。图 2 展现了这些数据存储器的总体数据结构。



注：为了简化图 2，这里只展示各结构的一个实例。在基准度量框架的实际实现中，每个结构可以界定多个实例。

图2 IT 项目绩效基准度量框架中的数据结构

IT 项目是基准度量框架中所关注的实体，具有各种属性以表示其性质或特性。

属性经测量（或记录），贮存在数据元素中。数据元素在 GB/T 41866 系列框架中是存取和操纵的最小数据单元（见 5.3）。“基准度量数据仓库”和/或“IT 项目数据”的拥有方为测量并维护数据而界定了数据元素。基准度量分析方也可以为其基准度量界定数据元素。

数据元素通过指定其性质来界定（见 5.6）。

注1：数据元素性质的示例有标识符、名称、数据类型和测量单位。

注2：在“IT 项目数据”中数据元素的情形中，数据元素界定的目的并不限于基准度量活动。数据元素也可以界定得用于促进项目管理、质量确保、财务管理等领域。

界定数据元素后，将界定一个或多个数据记录。数据记录是经界定的数据元素集，其中包括了表示 IT 项目所关注属性的所有必要的的数据元素。一个数据记录包含了特定 IT 项目的数据。在基准度量框架内，数据记录是维护并交换 IT 项目数据的典型对象。

界定数据元素和数据记录后，即能启动数据收集。通常情况下，IT 项目由其团队测量（或记录），并将数据记录提交到特定的验收台或存储到规定的数据存储设备中。收集到的数据记录可汇集并组织到几个可独立维护的 IT 项目数据集中。

注3：分类准则取决于数据拥有方和/或用户的需要（见 5.5）。

基准度量数据仓库中的数据 and/或“IT 项目数据”由基准度量分析方检索，以实施或发布基准。

## 5.2 属性

属性是实体中能以人工或自动化手段定量或定性区分出来的性质或特性。

IT 项目是 GB/T 41866 系列标准重要的关注实体。

注：在 GB/T 41866 系列中，IT 项目属性包括但不限于：

- IT 项目的输入（例如，要求、资源、可交付物）；
- IT 项目的结局（包括中间与临时产品，连同最终可交付物）；
- 关于 IT 项目本身的信息（例如，结构、过程、状况、进度）；
- 利益相关方（例如，角色、职责、专家经验）。

规划基准度量信息模型特定实例的第一步，通常是选出与基准度量需求有关的 IT 项目属性集。所选属性可采用数据元素形式测量、记录和维护，以支持日后各种信息需求。

### 5.3 数据元素

在 IT 项目绩效基准度量框架内，数据元素是：

- a) IT 项目中最小的数据单位；
- b) 通常以数据单元的形式在基准度量数据仓库和/或“IT 项目数据”中实现的；
- c) 由负责规划数据收集或基准度量的个人来界定的。

注1：该责任人员包括但不限于：基准度量数据仓库拥有方、IT 项目拥有方、基准度量分析方和基准度量用户。

为了支持基准度量活动，基准库和“IT 项目数据”通常包含有关 IT 项目说明性信息（例如，IT 项目名称及简要描述、IT 项目团队或数据提交方的联系详情）和/或其属性（例如，所适用测量方法的名称）。

注2：数据元素性质见附录 A。

### 5.4 数据记录

在 IT 项目中，因为有可能需要测量并维护以数十或数百计的数据元素，以满足利益相关方的多种信息要求，因此通常将数据元素结构化。

数据记录是有关数据元素经界定的组，其中包括所有必要的元素以表示关注的属性。一个数据记录对应一个特定的 IT 项目。该数据记录是测量、维护和交换基准度量活动中 IT 项目数据的单位。

注1：多重不同的数据记录有可能从某一组织中已界定的数据元素集合中来界定，以满足特别不同的信息要求，在这种情况下，一个数据元素有可能包括在多处数据记录中。

注2：数据记录往往以数据元素有序集的形式界定。这种数据记录的典型例子有电子表格中的数据表 and 数据库中的表。

示例：IT 项目组织界定 6 种数据元素（即项目名称、软件大小、总工作量、总工期、测试用例数，以及检测出的缺陷数），以此委任 IT 项目团队来进行测量和报告。该组织的质量保证部门界定过数据记录 DR1，其中包括项目名称、软件大小规模、测试用例数量和检测出的缺陷数。项目管理办公室也界定过数据记录 DR2，其中包括项目名称、软件大小、总工作量和总工期。一个 IT 项目团队测量 6 种数据元素，并以所要求的数据记录（即质量保证部门要求用 DR1、项目管理办公室要求用 DR2）形式提交数据。

### 5.5 IT 项目数据集

数据存储（基准度量数据仓库或“IT 项目数据”见图 C.1）中的数据记录，可以打包成组，以简化其用法和维护。这些组称为 IT 项目绩效基准度量框架内的 IT 项目数据集。

IT 项目数据集是带有类别的数据记录组，所包括的数据记录可按预先界定的准则选出。

注1：数据记录的分类准则有可能基于数据拥有方和/或用户的信息要求。数据集分类准则的例子有：

- 数据记录界定；
- 项目类型（例如，开发、增强和维护）；

- 业务和/或适用领域；
- 组织单位；
- 时段。

注2：数据记录有可能分类并存储到多处 IT 项目数据集。

注3：在实现物理数据存储时，多重 IT 项目数据集有可能会存入单一的物理存储器。

IT 项目数据集存储并维护在：

- 数据仓库拥有方（例如，基准供方或基准度量服务供方）；和/或
- IT 项目利益相关方（例如，IT 项目团队、IT 项目拥有方和基准度量用户）的 IT 项目数据（存储器）。

## 5.6 界定数据元素的要求

### 5.6.1 界定的性质综述

数据元素通过指定性质的集合来界定。表 1 中总结了强制的、建议的和可选的性质集合。也可指定补加性质以增强对数据元素的界定。

表 1 中的性质，见 5.6.2~5.6.16。

表1 界定数据元素的性质集

性质	优先	界定
〈一般性质〉		
标识符	[R]	规定数据元素的代码或序号
简称	[R]	规定数据元素的简称
名称	[M]	规定数据元素的名称
描述	[M]	通过指定数据元素的含义、目的等来规定数据元素的概况信息
优先级	[O]	规定数据元素的优先级（例如，强制的、建议的、可选的）
密级	[O]	规定披露的条件
〈范畴化性质〉		
可重现性	[M]	规定其值是否可重现（即“可重现”或“不可重现”）
可定标度性	[M]	规定其值是否可定标度（即“可定标度”或“不可定标度”）
〈与测量有关的性质〉		
数据类型	[M]	规定数据类型（例如，整数、实数、串、枚举）及其子性质
测量单位	[Ms]	规定测量单位[仅适用于可定标度的数据元素]
精度	[Rs]	规定值的所需精度[仅适用于数值数据类型]
时限	[R]	规定数据元素产生时间
规程	[M]	通过描述操作的逻辑序列（于适用时，连同算法一起），或者通过特定的现有规程表，规定如何生成数据元素的流程
源数据元素	[Rr]	规定产生值必须的数据元素表
依赖性	[M]	规定解释其值所必须的数据元素表（例如，限定符，见附录 A）
注 1：优先级（大写字母）：[M]——强制的，[R]——建议的，[O]——可选的。		
注 2：优先级（小写字母）：[s]——可定标度的数据元素，[r]——可重现的数据元素，〈无〉——适用于两者。		

### 5.6.2 标识符

数据元素宜有一个称为“标识符”的性质。标识符可以是一个代码或序号。标识符界定时，在 IT 项目数据集内所指应唯一。

注：标识符常在设计物理数据库或电子表格时予以引用。

### 5.6.3 简称

数据元素宜有一个称为“简称”的性质。简称界定时，所指在 IT 项目数据集内应唯一。

注：用户通常偏爱简称，而甚于技术上生成的标识符或全称，因为易记或易引用。

### 5.6.4 名称

数据元素应当有一个称为“名称”的性质。数据元素名称所指在 IT 项目数据集中应唯一。

### 5.6.5 描述

数据元素应当有一个称为“描述”的性质。

描述宜包括但不限于：

- a) 数据元素的意义；
- b) 测量（或记录）的目的；
- c) 于适用时，与其他数据元素的联系。

注：当意义、目的或联系从名称即显而易见时，描述就无需详尽。

### 5.6.6 优先级

数据元素可以有一个称为“优先级”的性质。

注：优先级定级实例是：“{强制的、建议的、可选的}”。

### 5.6.7 密级

数据元素可以有一个称为“密级”的性质。用于规定披露条件。

密级规范的实例有：

- 这个数据元素的值只应披露给基准度量数据仓库拥有方；
- 这个数据元素的值只应披露给质量保证部员工；
- 这个数据元素的值只应披露给预授权个人。

### 5.6.8 可重现性

数据元素应当有一个称为“可重现性”的性质。用于指明数据元素的值可否重现。

注：当且仅当源数据元素全包括在数据记录中时，某一数据元素的值才可重现。

### 5.6.9 可定标度性

数据元素应有一个称为“可定标度性”的性质。用于指明该数据元素的值是否具有测量标度（即“可测量的”或“描述性的”）。

注：描述性数据元素的示例子有 IT 项目名称、电话号码、文本解释、照片和电影。

### 5.6.10 数据类型

数据元素应有一个称为“数据类型”的性质。表 2 列出了数据类型的示例。

表2 典型数据类型及其范例子性质

数据类型	所界定子性质的示例
整数	取值范围（例如，按可接受的最小和最大数），显示格式（例如，数字逗号、加减号、消零、零值）
实数	取值范围（例如，按可接受的最小和最大数），记录精度（例如，单精度、双精度、四倍精度），显示格式（例如，数字逗号、加减号、消零、零值、指数部分）
货币	取值范围（例如，按可接受的最小和最大数），显示格式（例如，数字逗号、加减号、消零、币号）
文本串	最小和最大长度，字符编码集（例如，字母、字母数字、无限制、控制字符排除在外）
日期	输入格式（顺序、分隔符、消零、略去元素处理），显示格式（顺序、分隔符、消零、年份格式）
时间	输入格式要求（顺序、分隔符、消零、省略元素的处理），显示格式（顺序、分隔符、消零、小时格式）
枚举（从所给清单中选出）	替代项全表，多重选择可能性，编码和显示值及其格式
代码（输入预先界定的代码）	代码表访问信息（例如，URL、文件名），输入和显示格式（数码位数、分隔符、消零）

注：表2中代码的数据类型，是一种以系统化方式界定的字符序列。例如，电话号码或信用卡号、邮政编码、国家代码，或用户界定的代码。未定是可从预界定代码表中检索来输入值的。

#### 5.6.11 测量单位

数据元素有标度并存储数字值时，就应具有一个称为“测量单位”的性质。据元素接受多重测量单位之一时，应出示可接受单位清单。所选出和适用的单位应当记录在另一个数据元素。

注：对[工作量]，可接受单位清单示例：“{[人小时]、[人天]、[人月]、[人年]}”。

#### 5.6.12 精度

数据元素有标度并存储数值时，就应具有一个称为“精度”的性质。

注1：精度规范示例：“将值舍入到两位小数”。

注2：可重现数据元素的最大精度，通常取决于其源数据元素的精度。

数据元素的输入、显示和记录格式应满足这一精度规范。

注3：十进小数位数显示太多时，有可能对数据实际精度造成误解。

#### 5.6.13 时限

收集数据的用户需要 IT 项目的进展属性值时，数据元素应具有称为“时限”的性质，以此规定产生数据的可适用时间。

时限有时很重要，因为在 IT 项目生存周期中 IT 项目属性往往变更。当用户需要进行中的 IT 项目属性的快照值时，该性质应予准确规定。

示例1：项目经理希望在系统测试过程结束时实施实例，来检查[缺陷检出率]，以便制定决策将项目推进到下一阶段。为了满足基准度量要求，有必要为[检测出缺陷数]和[软件大小]规定时限，以确保两种测量结果表示同一开发阶段。

时限宜按国际标准标准化的软件开发生存周期来指定。引用特定组织的生存周期时，应映射到软件开发生存周期过程标准（见示例2）。

注：GB/T 8566和GB/T 22032界定了生存周期过程。

示例2：表3和表4列出了组织中特定生存周期模型和国家标准之间映射的示例。表3列出了国际软件基准度量标准化组织（ISBSG）的生存周期阶段到GB/T 8566—2022的映射，表4列出了信息技术促进机构（IPA）/软件质量增强中心（SEC）的生存周期阶段到GB/T 8566—2022的映射。

表3 特定组织生存周期模型和国家标准之间映射示例

序号	ISBSG 阶段	GB/T 8566—2022
1	规定	业务或使命分析过程
		利益相关方需要和需求定义过程
		系统/软件需求定义过程
2	设计	架构定义过程
		设计定义过程
		系统分析过程
3	构建	实现过程
		集成过程
4	测试	验证过程
5	实现	移交过程
		确认过程

表4 特定组织生存周期模型和国家标准之间映射示例

序号	IPA/SEC 阶段	GB/T 8566—2022
1	开发规划	业务或使命分析过程
2	需求界定	利益相关方需要和需求定义过程
		系统/软件需求定义过程
3	基础设计	架构定义过程
4	详细设计	设计定义过程
		系统分析过程
5	开发	实现过程
		集成过程
6	系统测试	验证过程
7	验收测试	移交过程
		确认过程
8	后续（操作）	运行过程
		维护过程

#### 5.6.14 规程

数据元素应当有一个称为“规程”的性质。规程是测量或记录数据元素值的逻辑运算序列（例如算法），可引用现有规程来进行规定。

注1：算法能以数学公式的形式来描述。

注2：引用现有规程的示例有：

- 适用国家标准功能规模测量方法之一；
- 用 {JPEG, JPEG2, GIF, TIFF, BMP} 之一给图片编码。

所规定规程产生多重结局（例如，同时有结果值和中间值）时，应明确指定目标值（见示例）。

示例：适用源代码检查器 XXX，并记录检测出的二级缺陷数。

当数据元素采用多重规程中之一来产生时，应呈送可接受的规程表，并将选出并适用的规程记录在另一数据元素。

### 5.6.15 源数据元素

可重现的数据元素宜有一种称为“源数据元素”的性质。以此规定产生数据元素值所必要的数据元素表。

注：源数据元素示例：引用数据元素[工作量]和[规模]，来产生数据元素[生产率]的值（其中，[生产率]=[工作量]/[规模]）。

### 5.6.16 依赖性

数据元素应有一个称为“依赖性”的性质。以此规定解释其值所必要的其他数据元素表（例如，限定符，见附录 A）。

## 5.7 界定数据记录的要求

数据存储器（基准度量数据仓库和“IT 项目数据”）拥有方应界定拟收集并维持的数据记录。

选择数据记录中的数据元素宜基于：

- 利益相关方的信息需求（例如，基准度量）；
- 数据产生的可行性和费效比；
- 可重现性和/或依赖性。

注1：GB/T 20917 界定了为满足特定信息要求而确定数据元素集合的过程。

注2：可重现数据元素的任何源数据元素不包括在数据记录中时，此数据元素将变得不可重现。如果可重现性丢失，必需修改对该数据元素的界定（例如，描述、规程、源数据元素）。

注3：如果一个数据元素依赖于其他数据元素，必需将后者一起包括在数据记录中。

注4：数据记录界定示例见参考文献[7, 10]。

如果一个组织旨在：

- 利用外部基准度量数据仓库和/或外部基准，或
- 向外部组织提交 IT 项目数据，

则应在选择并界定数据元素和数据记录时将兼容性考虑在内。

注5：为了使数据收集费用降低最低，可以将数据记录中数据元素的优先级分为下列 3 种范畴：

- [M]强制的，在所有 IT 项目中均强制收集和报告；
- [R]建议的，虽不带强制性，但强烈建议收集和报告；
- [O]可选的，收集和报告视情况逐一决定。

## 5.8 对界定 IT 项目数据集的要求

数据存储器（基准度量数据仓库和“IT 项目数据”）的拥有方宜界定 IT 项目数据集，以便能按下列两方面对不可预测的数据记录总量进行高质量管理：

- a) 对于数据拥有方的可维护性；
- b) 对于数据用户的可访问性和可用性。

注1：为数据拥有方将数据记录分类为多重 IT 项目数据集的示例有：

- 将数据记录分类为三组：数据验证之前、期间和之后；
- 将数据记录分类为两组：已完成的 IT 项目和进行中的 IT 项目。

注2：为数据用户将数据记录分类为多重 IT 项目数据集的示例有：

- 以组织单位划分数据记录；
- 以项目成绩划分数据记录（例如，做成的、质疑的和取消的）。

## 6 收集数据功能

### 6.1 概述

在 GB/T 41866 系列的基准度量框架中，数据收集主要通过三项活动进行：

- 测量 IT 项目；
- 数据提交；
- 维护数据仓库。

注：GB/T 41866.1 中界定了基准度量框架的活动（见图 C.1），GB/T 41866.2 中给出了每项活动的任务。

### 6.2 “测量 IT 项目”中的数据收集功能

在 GB/T 41866 系列的基准度量框架中，IT 项目测量的目的不限于将数据用于 IT 项目绩效基准度量。“测量 IT 项目”活动由若干 IT 项目利益相关方规划和实施，例如：

- 由 IT 项目团队或拥有方去了解和管理其项目进度；
- 由 IT 项目拥有方去了解并管理 IT 项目的绩效和进度；
- 由基准度量用户制备数据以实施；
- 由与 IT 项目有关的组织将数据提交给基准度量数据仓库拥有方。

“测量 IT 项目”活动中的数据收集功能，通常包括但不限于：

- a) 界定数据元素、数据记录 and IT 项目数据集的；
- b) 制备数据收集工具（例如，问卷、打印的数据录入表、计算机化在线/离线数据录入表）、数据验证工具和存储工具；
- c) 对“IT 项目数据”进行数据的征集、接收、验证和注册。

由于此项活动通常是在某一个组织或在关系密切的集体组织之内实施，因此组织的文化（例如，共用的数据元素、数据维护规程和测量方法）通常是相似的。

在这种情况下，如果 IT 项目数据限于在组织内使用，则不必十分详尽地指定数据元素的性总。可以引用预先界定的和共用的规则（例如，测量单位、精度、时限和规程）来界定数据元素。现有的工具和规程也可以适用于实施和支持“测量 IT 项目”活动。不过，当经测量的 IT 项目数据打算提交给外部组织或由外部组织使用时，就宜对数据元素严格界定。

### 6.3 “数据提交”中的数据收集功能

在 GB/T 41866 系列的基准度量框架中，提交 IT 项目数据的目的是提供 IT 项目数据以便 IT 项目绩效基准度量中注册和利用 IT 项目。

“数据提交”活动中的数据收集功能通常包括但不限于：

- a) 选择提交数据的 IT 项目；
- b) 制备拟提交的数据记录；
- c) 以数据记录或 IT 项目数据集的形式提交 IT 项目数据。

当需提交数据的数据元素或数据记录的界定并不够精准等效时，提交前就可能要有要追加像转换、规范化或重新测量之类的任务。

#### 6.4 “维护数据仓库”的数据收集功能

在“维护数据仓库”活动中，基准度量供方的职能通常包括但不限于：

- a) 界定数据元素、数据记录 and IT 项目数据集；
- b) 制备数据收集工具（例如，问卷、打印的数据录入表、计算机化在线/离线数据录入表）、数据验证工具和存储工具；
- c) 将征集、接收、验证和注册的 IT 项目数据存入数据仓库。

当 IT 项目数据是从文化不同的组织收集到时，数据元素的界定就宜尽可能详尽，以免引起数据提交时的误解和不必要的测量。当对同一数据元素有几种不同界定（例如，工作量有以人月、人日和人小时测量的）广泛跨组织使用时，就宜仔细考虑可转换性，以归并、整合出自多重组织的 IT 项目数据。

注：有时，将基准度量数据仓库称为“主数据仓库”，是由数据仓库拥有方拥有的存储器，由数据仓库管理方管理的仓库，用于存储和维护所收集的 IT 项目数据。“基准度量数据仓库产品”是主数据仓库的子集，将作为拟使用的产品交付。

### 7 数据收集的要求和指南

#### 7.1 数据源的要求和指南

众所周知，IT 项目的绩效在业务领域、软件领域、项目规模、开发类型（例如，新开发、增强、维护）、项目成败（例如，已成功、遭质疑、已撤销）等方面都存在差异。数据仓库拥有方宜缜密考虑逆收集数据的总量和质量，以满足基准与基准度量数据仓库产品目标用户的需求。

基准度量数据仓库拥有方希望更多地服务于受众，这使其乐意收集：

- 尽可能多的数据记录；
- 尽可能多的数据元素；
- 尽可能完备的高质量数据记录。

数据的体量、质量和完备性通常取决于数据收集的方式。数据收集的方式能分为三种类型：

- a) [选定的]，通过单独征集提交，从所选特定数据源（例如，组织、IT 项目团队）中收集数据；
- b) [公用的]，通过公共媒体征集提交从公众中收集数据；
- c) [委托的]，将数据提交委托给所规定 IT 项目团队或组织。

表5 数据收集类型的预期影响

数据类型	类型的优点	类型的缺点
[选定的]	能预期到数据质量高 易于控制数据的多样化	数据量有限
[公共的]	能预期到数据量大 能预期到数据完备性和质量上宽松多样化	经常会出现抽样偏倚（例如，只提交成功项目） 难以确保数据质量（例如，面对大量受众进行验证很难）
[授权的]	能预期到数据量大 能预期到目标受众全面覆盖	有可能纳入质量差的数据（例如，由于收集数据积极性不高）

注：抽样偏倚是系统误差，例如，由于从总体中样本的非随机性，导致总体中某些成员较其他成员纳入的可能性偏低。

确定收集数据的方式时，对优缺点（见表5）宜加以权衡。

## 7.2 时限指南

数据收集的时间安排能：

- a) 在特定时段征集和接收数据（例如，每年一次）；
- b) 全年一直接收数据。

在前一情况，数据仓库管理方（验证和注册）的工作量在一年内将严重失衡。因此，在此种情况下，有必要在短期内分配大量人力资源，同时宜详细界定流程和验证规则，以保持管理方工作的质量。

## 7.3 数据验证指南

数据收集职能（例如，“提交数据”和“维护数据仓库”两项活动的任务），包含下列几种有可能导致所收集数据无效的风险：

- a) 数据元素由提交方重新录入问卷或表单，在这一过程中有可能出现错误（例如，输入错误值、录入错误字段）；
- b) 提交方需要将其数据转换成符合要求的测量单位时，可能会出现错误（例如，计算或精度误差）；
- c) 提交方因缺乏信息而无法提供数据元素时，数据元素却有可能误作“0”（零）或“”（长度为零的串），而不是“N/A”（不适用）；
- d) 错误的源数据（例如，数据元素标记有误或数据总体出错）；
- e) 上游数据错误（例如，原始数据不正确）；
- f) 数据元素界定错误（例如，错误依赖性、错误算法）。

此清单上的前三种风险都是人为错误的例子，其他风险则是非人为错误的例子。

数据验证是检测此类错误的最佳最常用方法。数据仓库拥有方宜拟制一套验证规则，以确保验证过程和所得数据的质量。

典型的验证规则集包括但不限于：

- 检查数据类型，确认录入值以正确类型表示；
- 检查数据元素间的一致性（例如， $start\_date(开始日期) \leq end\_date(结束日期), total\_value(总值) = \sum detail\_value(各值求和)$ ）；
- 检查值的范围，所查值是否在合理范围内（例如，大小  $\geq 0$ ，百分值在  $[0, 100]$  范围内）；
- 通过比较值和现有数据的均值来检查值的范围（例如，现有 IT 项目的典型生产率值在  $[9.10, 23.4]$  范围内）。

注1：数据验证规则的示例见参考文献[8, 11]。

为了有效且高效地验证数据，数据仓库拥有方宜提请提交方通知联系点，以便数据仓库管理方进行质询。

作为验证的结果，数据的质量宜予评价并记录。

注2：数据质量指标例子有：

- A = 所提交数据已评估为完好，未认定出影响其完整性的任何问题；
- B = 提交数据看来基本完好，但有些因素可能会影响所提交数据的完整性；
- C = 由于未提供有重要意义的数据，因而无法评估所提交数据的完整性；
- D = 由于某一因素或某些因素，对所提交数据宜评为可信度很低。

## 7.4 隐私保护的要求和指南

保护数据提交方和其他所有需要隐私保护的利益相关方的隐私很重要，而且在许多情况下是不可或缺的。

注1：当 IT 项目正在开发客户特定的软件或系统，且项目团队已签署保密协定时，需将隐私保护的對象扩展到系统/软件的拥有方。

隐私保护方针应予编档并提供给数据提交方。这通常包括但不限于：

- a) 所提交数据的方式（例如，数据采用 U 盘、光盘等存储媒体或专用网络进行提交）；
- b) 所提交数据的管理过程（例如，将提交的数据存储在足够安全的计算机上，只能由得到授权的数据仓库管理方和分析方访问）；
- c) 所提交数据的用法限制（例如，数据将仅用于发行特定基准“XXX”，所提交的原始数据将永不披露）；
- d) 对有关隐私的信息的管理过程（例如，对此类信息的存储和管理要特别谨慎，重视匿名性和安全要求，仅供经授权人员访问）。

注2：与隐私有关的信息示例有提交方的姓名与联系信息、项目名称、系统名称、软件名称、对特定名称（开发支持工具的名称等）的引用。

注3：隐私保护方针文档的示例见参考文献[8]。

## 8 维护与支持职能

在 GB/T 41866 系列的基准度量框架中，对基准度量数据仓库和产品的维护与支持主要在两项活动中进行：

- “维护数据仓库”；
- “发布基准”。

这两项活动中的维护和支持职能包括但不限于：

- 维护基准度量数据仓库中的数据；
- 维护基准度量数据仓库本身；
- 交付基准度量数据仓库产品；
- 交付已发布的基准；
- 支持基准度量产品；
- 支持基准度量经验库。

注：GB/T 41866.1 中界定了基准度量框架的活动（见图 C.1），GB/T 41866.2 中引入了每项活动的任务。

## 9 维护与支持的要求和指南

### 9.1 数据维护的要求和指南

由于基准度量数据仓库中 IT 项目的数据记录通常存储和利用很长一段时间内，因此在某些时机有可能需要数据维护。

数据维护能按其简繁程度归入两种范畴：数据元素级维护和数据记录级维护。

数据元素级数据维护的常见起因有：

- a) 检测到数据值误差；
- b) 提交方请求添加/或收集数据值；
- c) 数据元素/数据记录界定的变更影响到了仓库现有数据；
- d) 组织范围的变更（例如，组织部门重组、软件生存周期过程变更）波及数据仓库现有数据。

注：组织变更的影响例如组织代码（或编码）更新。必需更新基准度量数据仓库现有数据记录中数据元素[组织代码]的值，需要替换成新码以确保现有数据记录连同未来新加数据记录的可访问性。

在这些情况下，基准度量数据仓库（即主数据仓库）的数据变更应正确地反映到每一产品（基准度量数据仓库产品、基准）。变更的影响严重时，宜将通知分发到基准度量数据仓库产品和基准的全体用户。

数据记录级数据维护的典型例子有：

- 过期的旧数据（例如，从注册起已超过 10 年）予以移除或分开管理；
- 巨大数据集分为多重数据集；
- 从数据仓库移除特定的数据记录；
- 界定新数据记录，并从现有数据记录创建初始数据集。

## 9.2 数据仓库维护的要求和指南

基准度量数据仓库出于与任何其他典型数据库相同的理由应予以维护。

数据仓库维护的范例操作有：

- 备份和恢复；
- 用户管理和访问控制（鉴别和授权）；
- 重组和重配置。

## 9.3 交付基准度量数据仓库产品的要求和指南

基准供方交付基准度量数据仓库产品。

一件基准度量数据仓库产品通常是基准度量数据仓库的一个子集。

注1：此处术语“子集”是指一件产品不必包括所有 IT 项目和/或来自“母”基准度量数据仓库的所有数据元素。

在此产生过程中，

- a) 所有与隐私有关的信息均应移除，并对存储媒体进行格式化处理；
- b) 应起草解释性报告。

注2：解释报告的要求和指南的界定见 GB/T 41866.3。

基准度量数据仓库产品应作为套件交付，通常包括但不限于：

- 数据产品，包括 IT 项目数据集；
- 解释性报告，描述数据元素界定、用户指南、联系信息等。

## 9.4 交付已发布基准的要求与指南

基准供方（例如，数据仓库拥有方）可将基准作为产品交付。

产品可以是：

- a) 从基准度量数据仓库分析及导出的一个或多个基准（例如，各种值）；
- b) IT 项目数据集，即基准度量数据仓库的子集。

在基准产品的产生过程中：

- c) IT 项目数据集应为基准度量数据仓库的子集；
- d) 应起草核心报告来描述已发布的基准，并应采用分析过程导出基准；
- e) 核心报告未包括必要信息时，应起草解释性报告。

注：对核心报告和解释性报告的要求和指南，在 GB/T 41866.3 中界定。

## 9.5 支持基准度量产品的要求和指南

基准度量产品（例如，基准度量数据仓库产品、已发布的基准）在交付后即开始使用，通常至少持续到下一个版本交付。因此，重要的是在交付后也要支持产品的使用。

产品交付后，对其用户的支持应包括下列功能：

- a) 回复对于产品的质询；
- b) 改正产品缺陷（例如，通过更换产品、发行勘误或修订）；
- c) 从用户处接收反馈，用以改进该活动各过程。

#### 9.6 支持基准度量经验库的要求和指南

为支持 GB/T 41866.2 中概述的“改进”任务组，宜对信息产品（例如，基准和基准度量报告）和基准度量的活动的评价连同得到的经验一起存入基准度量经验库。

注：GB/T 20917将“信息产品”界定为“解决信息需求的一项或多项指标及与其关联的解释”。例如，信息产品可以包括模板、图形、统计算法和解释指南。

仅限内部学习使用

附 录 A  
(资料性)  
基准度量信息模型

A.1 概述

基准度量信息模型是一种结构，将基准度量需求链接到关切的 IT 项目和所关注属性。其中描述：  
——在基准度量数据仓库和 IT 项目数据集中，对关切的 IT 项目属性如何测量（或记录）和维护；  
——对归档数据如何处理，以产生基准度量结局（例如，基准度量报告和基准），从而为利用基准度量结果提供基础。

该模型的总体结构见图 A.1。

注1：基准度量信息模型建立在 GB/T 20917 中界定的“测量信息模型”结构之上，用以指明信息需求和目标实体之间的联系。基准度量需求对应于信息需求，而 IT 项目的工程和业务对象对应于基准度量信息模型中的目标实体。

注2：图 A.1 中的模型远比 GB/T 20917—2007 附件 A 中的模型复杂。这仅仅是因为基准度量数据仓库或“IT 项目数据”通常包含数十或数百数据元素，以便能满足各种预期的基准度量需求，而 GB/T 20917 中的信息模型只处置一项或几项指标，并且仅满足一项信息需求。

选择或界定适当的数据以处置基准度量需求，始于一种观念，即 IT 项目中有哪些可测量（或可记录）属性与基准度量需求有关以及是何种关系。

规划数据收集的个人（例如，基准度量数据仓库拥有方、IT 项目拥有方，或基准度量分析方），来界定基准度量信息模型的组件，以此将 IT 项目属性链接到所规定基准度量结局来满足基准度量需求。基准度量信息模型有助于确定基准度量利益相关方在启动基准度量活动前需要规定哪些数据和规程。

基准度量信息模型由下列两种组件组构成：

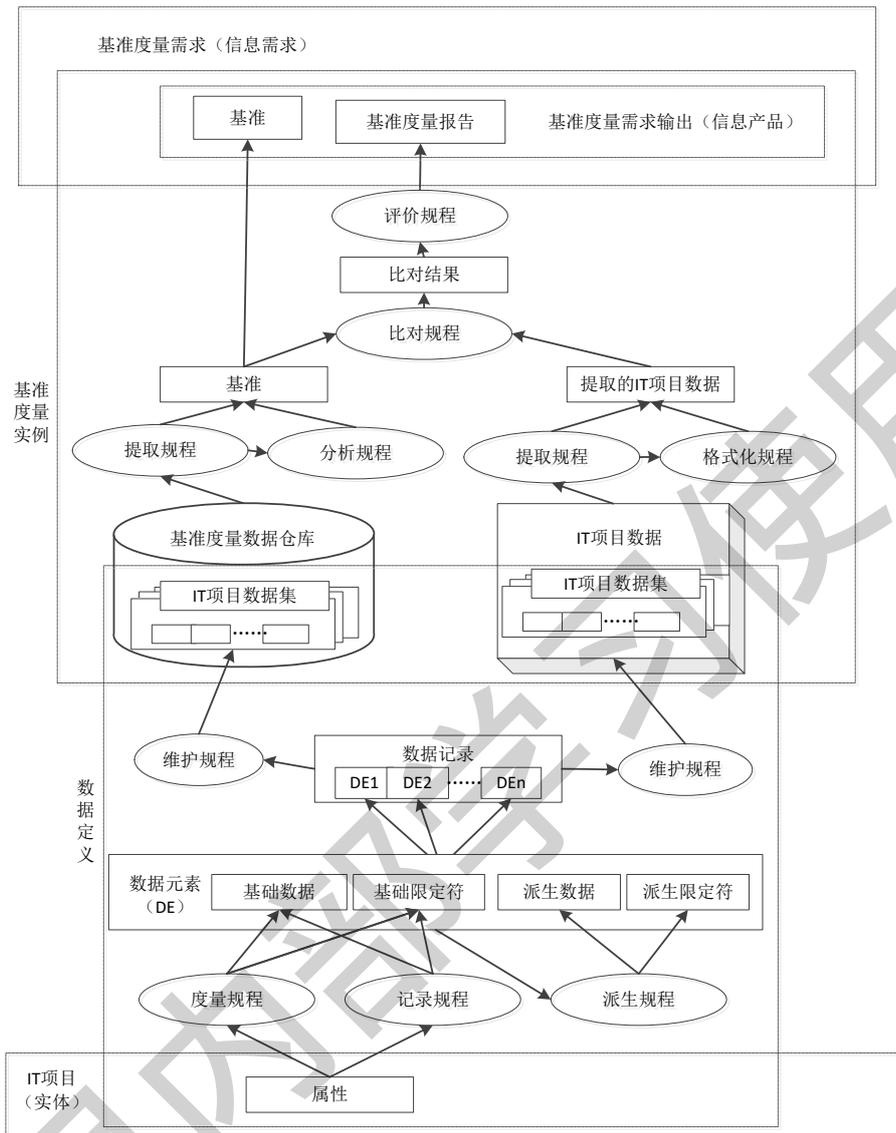
——“数据界定”组；

——“基准度量实例”组。

“数据界定”组包括各种数据组件（例如，数据元素、数据记录、IT 项目数据集）和操纵这些组件的各种规程组件。

注3：对基准度量数据仓库的界定通常是稳定的，以保持数据收集和服务的持续性。另一方面，对“IT 项目数据”的界定通常则更具动态性，因为 IT 项目团队有可能各有各的动态信息需求。

“基准度量实例”组有可能频繁重复“实施”活动。对这一组件组，图 A.1 中的模型指明了从输入数据（即基准度量数据仓库和“IT 项目数据”）到基准度量结局之间的粗略联系。这是因为此组件组的规程在基准度量的每一实例中都相当灵活地予以修订，以满足其临时性基准度量需要。



图A.1 基准度量信息模型

## A.2 模型组件

### A.2.1 IT 项目

GB/T 41866 系列的基准度量信息模型中，IT 项目是基准度量的对象，并且是重要的关注实体。

注：据GB/T 20917中的定义，实体是“种通过测量其属性来表特的对象”。

### A.2.2 属性

属性是某一实体中能以人工或自动化手段定量或定性区分出来的性质或特性。

注：该定义取自GB/T 20917—2007中3.2“属性”。

IT 项目是基准度量信息模型中重要的关注实体，具有多种属性来表示其性质或特性。

处置基准度量信息要求，有可能要求将分析范围大大拓宽，并有必要测量或记录多重性工程的和/或业务的属性。这些属性包括但不限于：

- 对 IT 项目的输入（例如，需求、资源、可交付物）；
- IT 项目的结局（包括中间和临时产品连同最终可交付物）；
- IT 项目本身（例如，结构、过程、调度表、状况、进度）；
- 利益相关方（例如，角色、职责、专家经验）。

通常情况下，规划基准度量信息模型特定实例的第一步，是选择最切合基准度量需要的 IT 项目属性集。选出的属性可以用多种方式测量或记录，并将其存储到多个数据元素，以支持未来不同的基准度量需要。

### A.2.3 数据元素

在 IT 项目绩效基准度量框架中，数据元素是：

- a) IT 项目中最小的数据单位；
- b) 通常作为基准度量数据仓库和/或“IT 项目数据”中数据单元来实现的；
- c) 由规划数据收集或基准度量计划的利益相关方来界定。

注1：利益相关方包括但不限于：基准度量数据仓库拥有方、IT 项目拥有方、基准度量分析方和基准度量用户。

从信息源的差异来看，数据元素划分成两组：

- 基础数据元素：通过将测量规程或记录规程适用于 IT 项目属性来测量或记录；
- 导出数据元素：通过将导出规程适用于一个或多个数据元素来产生。

注2：导出数据元素无需存入持久存储器，因为任何时候通过适用其导出规程都可重现。

为支持基准度量活动，基准度量数据仓库和“IT 项目数据”通常包含了有关 IT 项目的说明性信息（例如，IT 项目名称及简要描述、IT 项目团队或数据提交方的详细联系信息）和/或其属性（例如，所适用测量方法的名称）。这种信息通常以文本串的形式记录下来，没有测量标度。因此，从量标度的角度看，数据元素可以是下列两者之一：

- 有标度数据元素：是一种“测度”，并具有测量标度；
- 描述性数据元素：不具有测量标度（例如，姓名、电话号码）。

注3：测量标度类型的例子有：

- 标称的：值有范畴（例如，缺陷按其类型分类）；
- 序数的：值有等级顺序（例如，将缺陷指派为严重级）；
- 区间的：各值等距，对应于等量属性，但零点没有特定含义（例如，日历日期、摄氏度温度）；
- 比率的：测量值等距，对应于属性的等量，其中值 0（零）对应一个属性（例如，年龄、工作量、软件大小）都没有。

注4：描述性数据元素的例子有：[项目名称]、[项目描述]和[项目管理方名称]。

注5：描述性数据元素**不必是自由格式**。而会有某些记录规则（例如，名称长度限制到 20 个字符以内、只使用字母数字字符、电话号码需要从国家/地区编号开始并带前缀“+”）。

数据元素可以界定为能够保有不同标度之一的值（例如，工作量能以通过人小时、人天、人月之一来测量）。在这种情况下，有必要界定补加的数据元素，合而记录补充信息，以便解释这种数据元素的值。从这个角度看，数据元素可以是：

- a) 基数据：保有定量（即有标度）或定性（即描述性）值形式的 IT 项目属性的信息；
- b) 限定符：保有补充信息，以解释一个或多个数据元素（例如，测量方法、测量精度、测量单位、测量时限、标度换算因数）的值。限定符也可采用定量或定性值的形式。

示例：四种值 [工作量（单位：人年）]、[工作量（单位：人月）]、[工作量（单位：人天）]和[工作量（单位：人小时）]在逻辑上不同，因为是以不同标度来测量。如果这些值存储在不同数据单元，得到的数据表将具有四行[见图 A.2 a)]。该表实际效率很低，因为四个单元格中存值的只有一个，表格显得异常稀疏。引入个限定符[（工作量）单位]就解决了问题，将表格精化得高效[见图 A.2 b)]。此外，再引入个限定符[转换因子（到人小时）]，将能让用户通

过添加上导出数据元素[工作量（按人）小时]来比较带不同标度的值[见图 A.2 c) ]。

a)稀疏的表格				b)完善的表格 (1)		c)完善的表格 (2)			
[工作量] (年)	[工作量] (月)	[工作量] (日)	[工作量] (小时)	[工作量]	[单位]	[工作量]	[单位]	[转换因子]	[工作小时]
0.1				0.1	年	0.1	年	1680.0	168.0
	1.2			1.2	月	1.2	月	140.0	168.0
		24		24	日	24	日	7.0	168.0
			168	168	小时	168	小时	1.0	168.0

↑ 限定符 (1)                      ↑ 限定符 (1)                      ↑ 限定符 (2) = [工作量] × [转换因子]

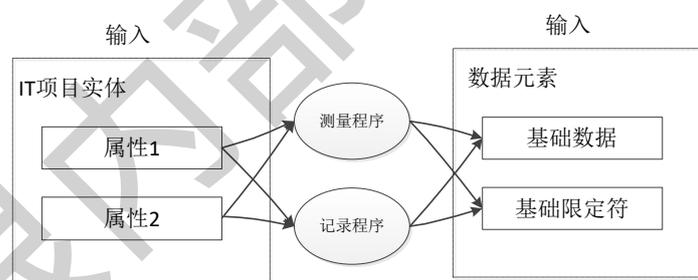
图A.2 限定符元素的范例用法

#### A.2.4 测量规程

测量规程适用于测量 IT 项目的属性，以产生有标度的基础基数据。测量规程也可产生基础限定符来记录补充信息，以此解释特定基数据的值。复杂的测量规程也可产生描述性基数据。测量规程可能的输入输出请参见图 A.3。

测量规程是操作的一种逻辑运算序列，用于给特定属性定量和/或定性。这些运算通常包括但不限于对“出现”计数或观察时间推移之类的任务。

同一测量规程可适用于多个属性，以测量多重基础数据元素。



图A.3 测量规程和记录规程的输入输出

#### A.2.5 记录规程

记录规程适用于记录描述性的基础基数据和基础限定符。

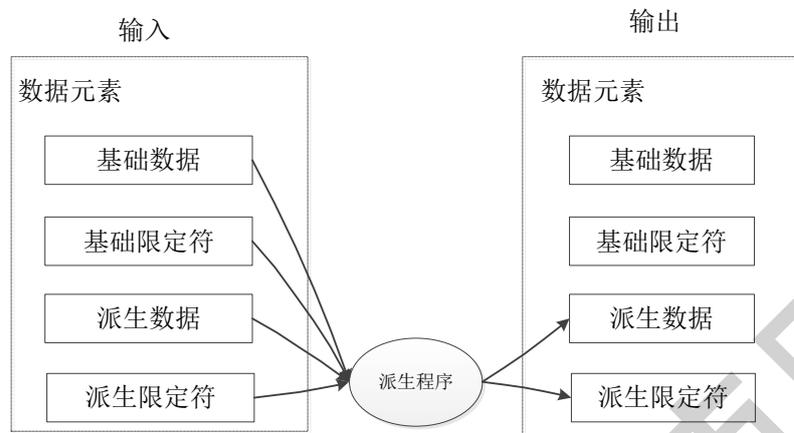
记录规程是一种逻辑运算序列，用于将 IT 项目属性的特性记录到描述性基础基数据。这些运算可包括各种活动，但不限于打印文本、拍照、绘图、记录影像或音响。同一记录规程可适用于多个属性，以记录多重描述性基础数据元素。

记录规程也可产生基础限定符，以记录补充信息，来解释特定数据元素的值。

#### A.2.6 导出规程

导出规程是一种逻辑运算序列，适用于产生导出的元素。

任何数据元素都可作为导出规程的输入。导出规程的输出则始终是导出数据元素（见图 A.4）。



图A.4 导出规程的输入输出

### A.2.7 数据记录

数据记录是一种经界定的有关数据元素的集合，其中包括了表示 IT 项目关注属性所必须的全部数据元素。一个数据记录对应一个特定的 IT 项目。数据记录是测量、维护、交换基准度量活动中 IT 项目数据的单位。

注1：多个不同的数据记录，有可能从某一组织已界定数据元素集合中来界定，以满足特别不同的信息需要。在这种情况下，一个数据元素有可能包括在多重数据记录中。

注2：数据记录常以数据元素有序集的形式来界定。这种数据记录的典型例子有电子表格软件里的数据表和数据库中的表。

### A.2.8 IT 项目数据集

由于基准度量数据仓库和 IT 项目数据通常存有储大批 IT 项目的数据，所以可将数据记录归拢成组，使其易于使用和维护。

IT 项目数据集是一种分好类的数据记录组，其中所收集的数据记录由预先界定的准则选出。

注1：分类准则取决于数据的拥有方和/或用户的信息需求。分类准则的例子都是将 IT 项目数据范畴化，通过：

- 数据记录界定；
- 项目类型（例如，开发、增强和维护）；
- 业务和/或应用领域；
- 组织单位或 IT 项目团队；
- 选择特定时段的 IT 项目。

注2：数据记录有可能加以分类并存储到多重 IT 项目数据集。

注3：在实现数据物理存储器时，多重 IT 项目数据集有可能存储到单一物理存储器中，并通过附上限定符来区分数据记录所属的数据集。

IT 项目数据集存储和维护在：

- 数据仓库拥有方的基准度量数据仓库（例如，基准供方和基准度量服务供方）；
- IT 项目利益相关方（例如，IT 项目团队、IT 项目拥有方和基准度量用户）的“IT 项目数据”。

### A.2.9 维护规程

维护规程注册并维护 IT 项目数据集的数据记录。维护规程的任务包括但不限于：

- 添加、更新、删除 IT 项目数据集中的数据记录；
- 备份和/或恢复 IT 项目数据集中的数据记录；
- 创建和/或修改对数据元素或数据记录的界定；
- 优化数据物理存储器；
- 将数据记录移送或拷贝到其他 IT 项目数据集。

#### A. 2.10 提取规程

提取规程从基准度量数据仓库或“IT 项目数据”中选出 IT 项目数据集的一个子集。在提取过程中，创建一个子集是通过：

- a) 忽略数据元素组中不必要的数据元素；和/或
- b) 只选出符合选择准则的特定关注 IT 项目。

提取规程的结局有可能是一个基准或所提取 IT 项目数据集。在某些情况下，提供基准需要特别的分析，而提供在 IT 项目数据集中已提出的数据需要规格化。

注：在 GB/T 41866 系列的框架中，基准和所提取 IT 项目数据均有可能是：

- 数值（例如，数据元素的值、统计量）；
- 选出的 IT 项目的 IT 项目数据集；
- 统计图表（例如，散点图、直方图）及其他。

#### A. 2.11 分析规程

分析规程是一种逻辑运算序列，用以导出一个基准。

注：统计分析技术通应适用于导出各种基准。

#### A. 2.12 规格化规程

规格化规程是一种逻辑运算序列，用以使数据可比较。

注1：将数据规格化的原因，举例有：

- 不同的数据界定；
- 丢失的数据值；
- 不同的计量单位；
- 不一致的数据汇集。

注2：规格化的例子有：

- 转换数据元素的测量单位和/或值；
- 对 IT 项目的数据集补加一个非特征式导出数据元素（于各输入数据元素都可用时）；
- 补加一个基础元素，并通过某一逻辑或数学函数（例如，集结、分割、内插）。

#### A. 2.13 比对规程

比对规程是一种逻辑运算序列，用以将所提取 IT 项目数据与基准进行比对。

比对基于一组假设和基准与所提取 IT 项目数据集之间的预期联系。该规程提供了一种观察和调查结果的汇集，预期关切到所指基准度量需求。典型的比对结果都以指标、统计图表、文本描述等形式来表示。

#### A. 2.14 评价规程

评价规程依照预先界定的判断准则，审查比对结果满足基准度量需要的适用性。评价规程有可能由基准度量分析方实施和/或自动化实施。

注：在GB/T 41866系列中，判断方法和准则需要在“实施”活动的规划阶段来界定。（见GB/T 41866.2）。

比对结果判定为满足判断准则时，基准度量分析方即产生一份基准度量报告，并将其提交给基准度量用户。

仅限内部学习使用

## 附录 B

(资料性)

## 数据元素分类和与其他有关标准的比较

## B.1 概述

GB/T 41866 系列中基准度量信息模型的数据元素，本附录从测量目标、产生过程和测量标度级别三个方面分为八种范畴。

本附录还将各范畴的数据元素的特征与 GB/T 20917 中的测度或 GB/T 25000.21 中的质量测度元素 (QME) 进行了比较。表 B.1 概括了三项国标之间的全面差异。

表B.1 数据元素特性与有关标准的比较

	GB/T 41866 系列中的数据元素	GB/T 20917 中的测度	GB/T 25000.21 中的质量测度和 QME
界定	数据元素： a) IT 项目数据最小的单位； b) 在基准度量数据仓库或 IT 项目数据中，通常作为数据单元来实现	测度： 作为测量结果赋予某一值的变量	质量测度量： 导出测度，界定为质量测度元素的两个或更多值的一种测量函数。 质量测度元素 (QME)： 依性质及其量化的测量方法界定的测度，其中可选地包括通过一种数学函数所作的变换
原始性质	实际的 (物理的)	逻辑的 (理论的)	逻辑的 (理论的)
变化量	基础基数据 导出基数据 基础限定符 导出导出限定符	基础测度 导出测度 指标	质量测度 质量测度元素
无标度的描述性数据	作为描述性数据元素包括在范围内	在范围外	在范围外
限定符	作为限定符包括在范围内	提及其必要性	未提及
配置	实例化为基准度量数据仓库或 IT 项目数据中一个数据单位，用来存储值	界定为逻辑变量，但未实例化	界定为逻辑变量，但未实例化
数据结构	IT 项目数据集由表示某一 IT 项目的数据元素集所组成	无界定的数据结构	无界定的数据结构
数据管理	IT 项目数据集的数据元素集	未提及	未提及
数据类型	数据类型和精度以显式界定	未显式提及	
对数据界定的要求	详细指定为数据元素界定要求	仅在资料性附录中给出范例性界定	详细规定为 QME 界定要求

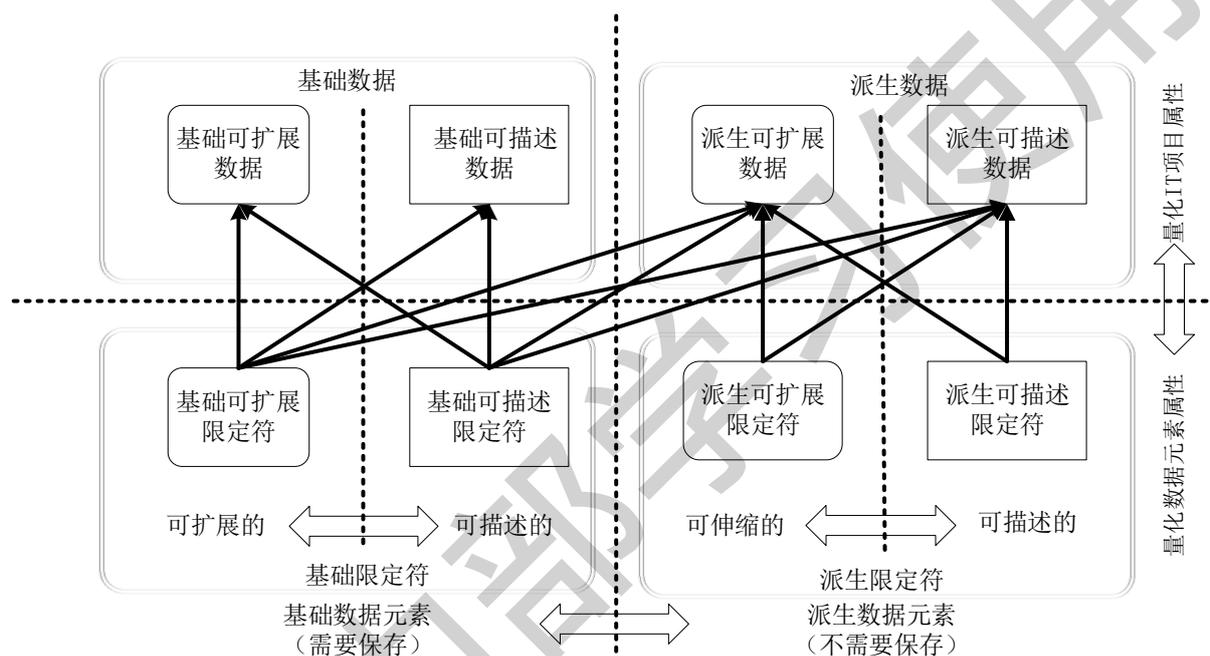
## B.2 数据元素分类

### B.2.1 总则

GB/T 41866 中的数据元素能按下列三个方面进行分类：

- 测量（或记录）的目标；
- 产生过程；
- 测量标度的级别。

由于这三个方面都能适用于每一数据元素，所以就有八种（= 2×2×2）数据元素范畴，见图 B.1。



图B.1 GB/T 41866 系列中数据元素的分类

#### B.2.1.1 测量（或记录）的目标

数据元素分为：

- 基础数据：用来量化（或记录）IT 项目属性的数据元素；
- 限定符：用于量化（或记录）解释一个或多个数据元素值的补充信息。

注1：限定符与其他数据元素一起予以测量、记录或导出（通过测量方法、记录方法或导出规程）。

注2：基础数据的限定符示例[接受测试过程的工作量]包括：

- [工作量单位]，当工作量数据由数据供方内多个单位测量时，这种工作量单位对于存档数据就必不可少；
- [转换因子（每月小时数）]，对于数据用户将工作量数据规范化为想要的标度或单位是必不可少的。

注3：限定符可以限定多个数据元素。 这种情况的一个示例是：限定符：[工作量单位]限定各数据元素[XXXX 过程的工作量]。

#### B.2.1.2 产生过程

GB/T 41866 中的数据元素也能分为：

- 基础数据元素：通过将测量（或记录）方法适用于 IT 项目属性所测量（或记录）的数据元素；
- 导出数据元素：以导出过程形式来界定的数据元素。

这种分类对于基数据和限定符均可适用。

注1：导出的数据元素任何时候都通过适用导出规程都可重现，因此不必将其存入持久性存储器。

注2：GB/T 41866 系列对“指标”与所导出数据元素不加区分。术语“指标”在图 A.2 中并未出现，因为基准度量的结果通常不能用单一数值来解释。

### B.2.1.3 测量标度的级别

数据元素也可分类为（见图 B.2）：

- 有标度数据元素：具有标度的数据元素；
- 描述性数据元素：不具有标度的数据元素。

这种分类同样适用于基数据和限定符。

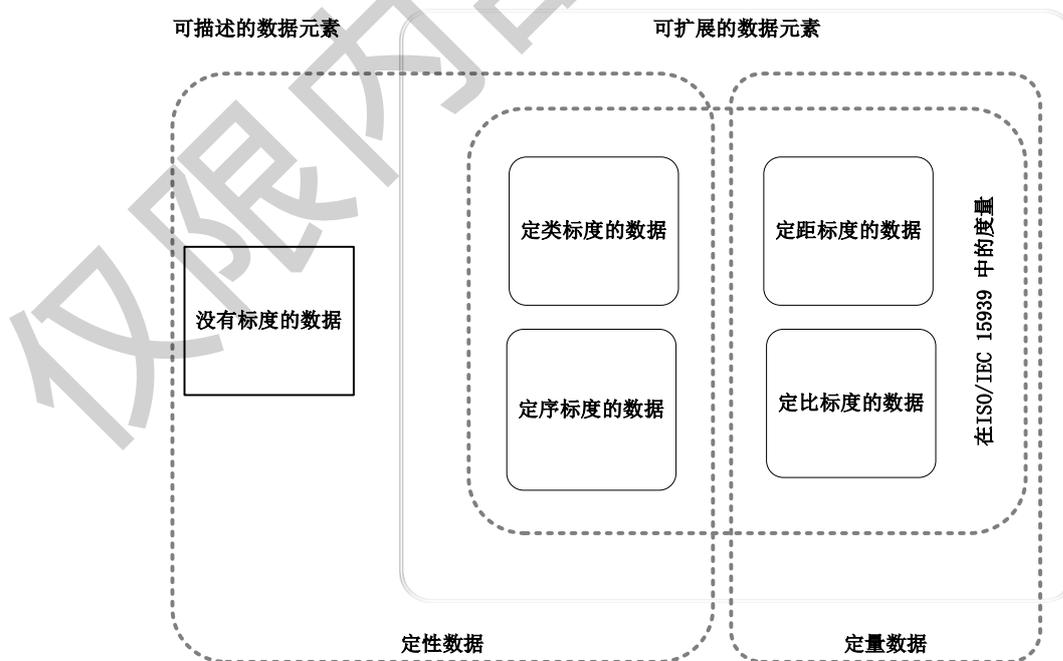
有尺度的数据元素再分为下列子类。

- “定类尺度的”，测量值是分类的。例如，按类型对软件缺陷进行分类，不考虑类别之间的顺序。
- “定序尺度的”，测量值是分级有序的。例如，将软件缺陷按严重级别排了序。
- “定距尺度的”，测量值等距，对应于同等数量的属性。例如，圈复杂度的最小值为 1，但每个增量表示一条附加路径，值 0 是不可能的，这个数据最小距离是 1。
- “定比尺度的”，测量值等距，对应于同等数量的属性，其中值 0 与任何属性都不对应。例如，以数量表示的大小需求是一个比率尺度，因为值 0 对应于没有需求，而补加界定的每一需求都表示相等的增量数量。

注1：描述性数据是基准度量中必不可少的信息，用来记录数据的解释性信息（例如，IT 项目的简档信息、IT 项目的标识或数据提交方的联系方式）。

注2：GB/T 41866 系列避免使用术语“测度”（不论用作动词，还是用作名词）来描述数据元素。描述性数据元素并不是测量理论中的“测度”（名词），只是“所记录的”而非“所测量的”。

注3：描述性数据超出 GB/T 20917 和 GB/T 25000.21 的范围。



图B.2 数据元素的测量标度的级别

### B.2.2 基础有标度基数据

基础有标度基数据是下列数据元素：

- [测量目标]，量化 IT 项目属性；
- [产生过程]，通过人工或自动手段适用特定的测量方法来测量；
- [测量标度级]，为有标度值。

注：在GB/T 20917—2007的3.3中，这种类型的基础数据元素归类为“基础测度”。

### B.2.3 基础描述性基数据

基础描述性基数据是下列数据元素：

- [测量目标]，记录 IT 项目属性；
- [产生过程]，通过人工或自动手段适用特定的记录方法来记录；
- [测量标度级]，为无标度值。

注1：基础描述性基数据的示例有[项目名称]、[项目描述]和[领导者名称]。

注2：此类型的基础数据元素，不是GB/T 20917—2007的3.3中的“基础测度”，因为没有标度。

### B.2.4 基础有标度限定符

基础有标度的限定符是下列数据元素：

- [测量目标]，保留了补充信息以解释特定数据元素的值；
- [产生过程]，通过适用特定的测量方法，连同基础数据元素，以人工或自动手段进行量化；
- [测量标度级]，为所定标度的值。

注1：基础有标度限定符最初只限定一个或多个基础数据元素。当所限定基数据元素的环境信息接替导出元素时，限定联系可以继承到导出的数据元素。

注2：基础有标度限定符的一个例子是转换因子[工作小时每天]将[工作量]从[人天]转换为[人小时]。

注3：这一类型的基础限定符归类到GB/T 20917—2007的3.3中的“基础测度”。

### B.2.5 基础描述性限定符

基础描述性限定符是下列数据元素：

- [测量目标]，保留了补充信息以解释特定数据元素的值；
- [产生过程]，通过适用特定的记录方法，连同基础数据元素，以人工或者自动的手段记录下来；
- [测量标度级]，为无标度值。

注1：基础描述性限定符最初仅限定一个或多个基础数据元素。当限定基础数据元素的环境信息继承到导出元素时，限定关系可以继承到导出数据元素。

注2：基础描述性限定符元素的一个示例是自由格式的[所适用计数工具的名称]。当备选工具有限时，该数据元素能界定为枚举式清单（即标称标度）中的选项。

注3：这一类型的基础限定符，不是GB/T 20917—2007的3.3中的“基础测度”，因为没有标度。

### B.2.6 导出的有标度基数据

导出的有标度基数据是下列数据元素：

- [测量目标]，将 IT 项目属性量化；
- [产生过程]，通过适用导出过程（连同从中引用或接替环境信息）从一个或多个数据元素（任何种类）导出；
- [测量标度级]，为所定标度值。

注1：在GB/T 41866系列中，一个数据元素（任何种类）的变换视为新导出的有标度基数据，因为这一转换（例如，变更测量单位）在基准度量中给出有效值。

根据 GB/T 20917 和 GB/T 25000.21 中的界定，导出的有标度基数据能分为下列三种类型：

——[类型-1] 由两个或更多基础有标度基数据产生的导出有标度基数据；

注2：该类型在 GB/T 20917 中归类到“导出测度”，在 GB/T 25000.21 中归类为“质量测度”。

注3：这一类型的例子是： $[\text{缺陷密度率}] = [\text{检测到的缺陷数}] / [\text{软件大小}]$ 。

——[类型-2] 仅由一个基础有标度基数据产生的导出有标度基数据；

注4：这一类型的一个示例是数据元素单位的变换： $[\text{工作量}]$ 从 $[\text{人小时}]$ 到 $[\text{人月}]$ 。

注5：在 GB/T 20917 中，此类型未归类为“导出测度”。GB/T 20917—2007 的 A.2.4 中规定：“基础测度的简单变换（例如，取基础测度的平方根）并不添加信息，因而不产生导出的测度”。另一方面，在 GB/T 25000.21—2019 中 4.14 的定义中，则视为新“质量测度”（导出测的）。

——[类型-3] 仅由一个或由多个基础的或导出的描述性数据元素产生的导出有标度基数据。

注6：这一类型的示例是一种导出规程，用于测量和存储 $[\text{项目名称}]$ （包含有文本串）的“字符数”。 $[\text{项目名称}]$ 值是无标度的，在首次使用导出规程时测出。

注7：在 GB/T 20917 和 GB/T 25000.21 中，此类型均归类到“基础测度”，因为其能被解释为“针对 IT 项目属性的测量是在当时通过导出规程实施的”。

### B.2.7 导出的描述性基数据

导出的描述性基数据是一种数据元素，用以存储将导出规程适用于任何数据元素所产生的无标度值。

导出的描述性数据元素是下列三种数据元素：

——[测量目标]，记录 IT 项目属性的数据元素；

——[产生过程]，通过导出规程从一个或多个（基础的或描述性）数据元素中导出；

——[测量标度级]，为无标度的值。

注1：这种类型的示例有：对基础或导出的描述性数据元素的文本运算（例如拼接、重新格式化、变更字符码）。

注2：这一类型的导出数据元素，并不是 GB/T 20917 中的“测度”，因为值没有标度。

### B.2.8 导出的有标度限定符

导出的有标度限定符是下列数据元素：

——[测量目标]，用于记录补充信息以解释特定导出数据元素的值；

——[产生过程]，通过适用导出规程，连同从任何范畴的一个或多个数据元素所导出的数据元素产生的；

——[测量标度级]，为定标度的值。

通过将导出规程适用于一个或多个（基础的或描述性）数据元素所导出的

根据 GB/T 20917 和 GB/T 25000.21 中的定义，导出的有标度限定符能分为以下两种类型：

——[类型-1]，通过适用导出规程与据类型-1 导出的有标度基数一起产生，以保留补充信息去解释特定导出数据元素的值。

注1：在 GB/T 20917—2007，3.3 中，这种导出的限定符被归类为“导出测度”。

——[类型-2]，通过适用导出规程与类型-2 或类型-3 导出的有标度基数据一起产生，以保留补充信息去解释特定导出基数据的值。

注2：在 GB/T 20917—2007 的 3.3 中，这一类型的导出数据元素归类为“基础测度”。

### B.2.9 导出的描述性限定符

导出的描述性限定符是下列数据元素：

——[测量目标]，记录补充信息以解释特定基础的或导出的基数据的值；

——[产生过程],通过应用导出规程与来自任何范畴的一个或多个数据元素的任何导出数据元素一起产生;

——[测量标度级],为无标度值。

注:这一类型的示例是对基础的或导出的描述性限定符数据元素的文本运算(例如,拼接、重新格式化、变更字符码)。

注:该类型的基础限定符,并不是GB/T 20917-2007的3.3中的“基础测度”,因为没有标度。

仅限内部学习使用

附录 C  
(资料性)

GB/T 41866.1-2022 中 IT 项目绩效基准度量框架模型

本附录概述了 IT 项目绩效基准的总体框架（见图 C.1）。

注：有关基准度量框架及其概念和界定的更多信息，见GB/T 41866.1。

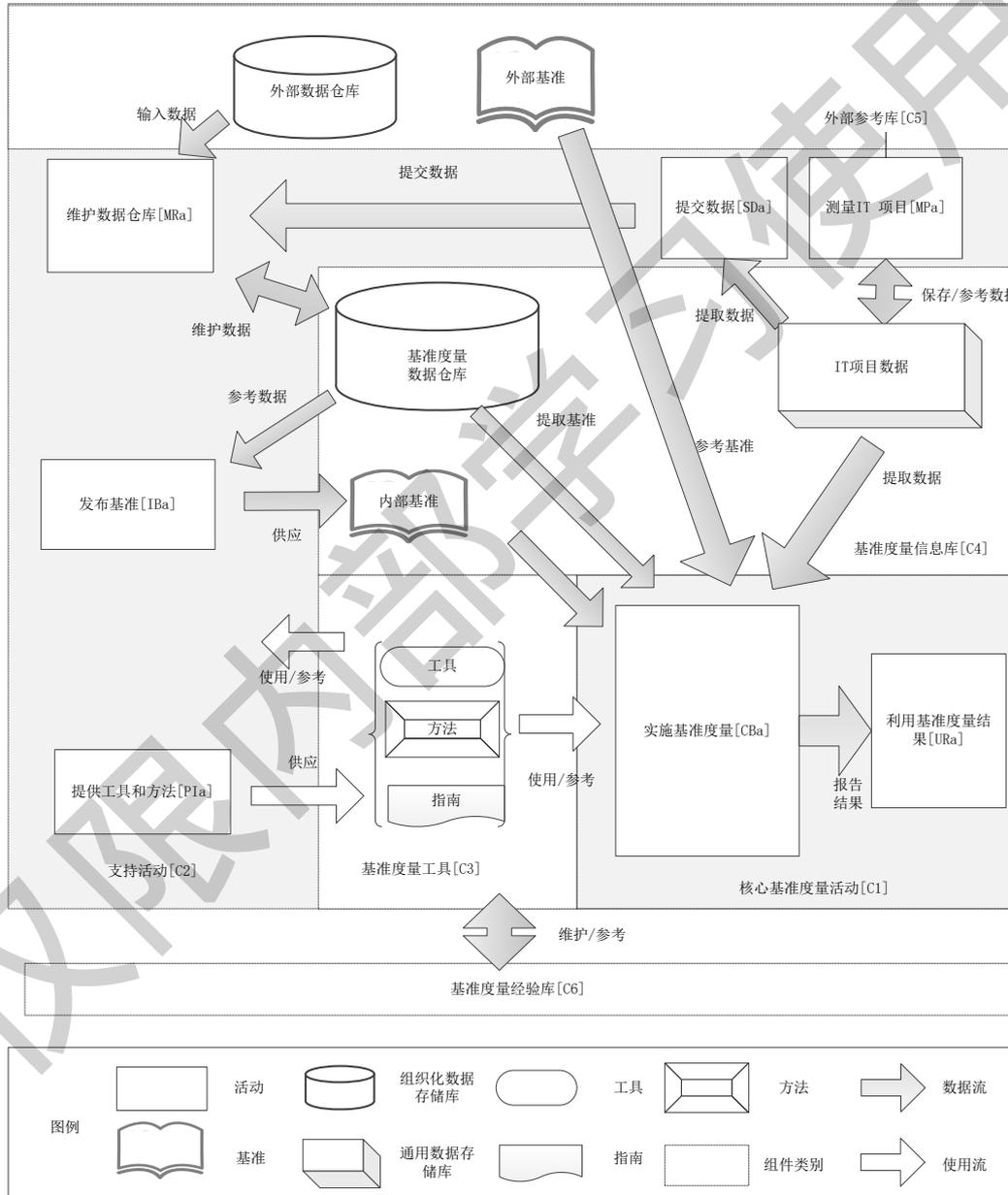


图 C.1 IT 项目绩效基准度量框架

该框架由下列 6 种范畴的活动或组件构成：

- [C1] 各种“核心基准度量活动”，共同说明了基准度量的某一实例；
- [C2] 各种“支持活动”，提供基准度量信息库以及用于基准度量实例的工具与方法；
- [C3] 各种“基准度量工具与方法”：由利益相关方用来实施基准度量实例；

- [C4] “基准度量信息库”，包含基准度量实例期间使用的数据；
  - [C5] “外部参考库”，为基准度量实例提供了备选的或补加的外部数据（即数据仓库和/或基准）；
  - [C6] “基准度量经验库”，包括了从当前和/或以往基准度量经验中获得的知识和经验。
- 前两种范畴（C1 和 C2）引入了用于主导或支持基准度量实例的活动，包括 7 项活动：
- “实施基准度量”活动（CBa），启动并执行一个基准度量实例；
  - “利用基准度量结果”活动（URa），将基准度量结果应用于各种商务目的；
  - “测量 IT 项目”活动（MPa），测量 IT 项目制品并维护数据；
  - “提交数据”活动（SDa），选择并提供 IT 项目数据，将其包含到基准度量数据仓库中；
  - “维护数据仓库”活动（MRa），接收、验证并存储 IT 项目数据存储到基准度量数据仓库中，日后对该库进行管理；
  - “发布基准”活动（IBa），分析基准度量数据仓库中的 IT 项目数据，并提供内部基准；
  - “提供工具与方法”活动（PIa），开发并提供工具、方法和指南，以支持基准度量框架中各项活动。

为有效的基准度量，上述 7 项活动均需予以正确执行。

第 3 种范畴（C3）界定了三种类型的基准工具与方法：

- 工具，提供支持利益相关方的手段；
- 方法，提供规程；
- 指南，提供指导性信息。

其余 3 种范畴（C4、C5 和 C6）由各种类型的信息库组成：

- 基准度量数据仓库，是一种数据存储器，为 IT 项目维护可靠数据，用于产生基准；
- IT 项目数据，维护与各种 IT 项目有关的数据；
- 内部基准，提供一组可共同使用并经授权的预先确定的基准；
- 外部数据仓库，一种由外部组织维护的基准度量数据仓库；
- 外部基准，通过分析外部数据仓库而在某些外部组织中发布；
- 基准度量经验库，关于基准度量结果和所得经验的档案库。

GB/T 41866 系列中“内部”“外部”的概念，均出自利益相关方的视角，即取决于利益相关方的判断。例如，当甲组织决定采用乙组织的 IT 项目数据时，乙组织的数据就视为甲组织内部数据的一部分。这种采用组织外的数据或基准的情况，经常发生在集团企业之内或关系密切的业务伙伴之间。

## 参 考 文 献

- [1] GB/T 8566—2022 系统与软件工程 软件生存周期过程
- [2] GB/T 20917—2007 软件工程 软件测量过程
- [3] GB/T 22032 系统与软件工程 系统生存周期过程
- [4] 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
- [5] GB/T 25000.21—2019 系统与软件工程 系统与软件质量要求和评价 (SQuaRE) 第21部分: 质量测度元素
- [6] GB/T 36964—2018 软件工程 软件开发成本测量规范
- [7] IPA/SEC. IPA/SEC data item definitions, version 3.0 (for IPA/SEC White Paper 2007 on software development projects in Japan), Software Engineering Center, Information-Technology Promotion Agency, Japan (IPA/SEC), freely available at <http://www.ipa.go.jp/english/sec/>
- [8] IPA/SEC. IPA/SEC data item validation rules, version 3.0 (for IPA/SEC data item definitions, version 3.0), Software Engineering Center, Information-Technology Promotion Agency, Japan (IPA/SEC), freely available at <http://www.ipa.go.jp/english/sec/>
- [9] ISBSG. Confidentiality and integrity of the data, International Software Benchmarking Standards Group (ISBSG), <http://www.isbsg.org/>
- [10] ISBSG. Field Descriptions: Data Release 11, International Software Benchmarking Standards Group (ISBSG), freely available at <http://www.isbsg.org/>
- [11] ISBSG. Data validation standard, International Software Benchmarking Standards Group (ISBSG), freely available at <http://www.isbsg.org/>
- [12] ISO/IEC 29155-1, Systems and software engineering—Information technology project performance benchmarking framework—Part 1: Concepts and definitions
- [13] ISO/IEC 29155-2, Systems and software engineering—Information technology project performance benchmarking framework—Part 2: Requirements for benchmarking