INTERNATIONAL STANDARD

ISO/IEC 29155-4

First edition
2016-10-15

Systems and software engineering — Information technology project performance benchmarking framework —

Part 4: Guidance for data collection and maintenance

Ingénierie des systèmes et du logiciel — Cadre de conduite de tests de performance de projet de technologies de l'information —

Partie 4: Directives pour la collecte de données et la maintenance

Reference number
ISO/IEC 29155-4:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is ISO/IEC JTC1, *Information technology*, Subcommittee SC 7, *Software and systems engineering*.

A list of all parts in the ISO/IEC 29155 series can be found on the ISO website.

Introduction

Benchmarking is an activity of comparing “objects of interest” to each other or against a benchmark to evaluate characteristic(s). In the context of the ISO/IEC 29155 series, the “object of interest” is the performance of information technology (IT) project, and the characteristic is a particular aspect of an IT project such as productivity.

Benchmarking is one of the fastest-growing techniques in the area of IT project management. Instances of IT project performance benchmarking are initiated and conducted for various reasons. Among the most common reasons are the following:

a) the need to compare project productivity between similar industries;

b) the need to compare productivity between different project types and technologies;

c) the need to find the most effective targets for IT development process improvement;

d) the need to compare productivity between different suppliers;

e) the need to improve project management maturity;

f) the need to improve project estimation capability.

Much has been written regarding the trials of establishing IT project performance benchmarking, and statistics bear witness to the high failure rate of measurement and benchmarking programs. The most probable causes for failure have been disappointment in the benchmarking outcomes due to a lack of alignment between the selected measurements and business goals, and the misunderstanding of project level measurements in relation to program and portfolio management levels. When there is no alignment between executed measurements and provided outcomes, unnecessary effort is required from the IT project teams collecting the project data. The result is decreased motivation to continue and institutionalize benchmarking.

As is shown in Figure 1, the ISO/IEC 29155 series consists of multiple parts.

— ISO/IEC 29155-1 provides the overall framework model for IT project performance benchmarking. It consists of activities and components that are necessary to successfully identify, define, select, apply, and improve benchmarking.

— ISO/IEC 29155-2 prescribes the required tasks in individual benchmarking activities that are necessary to execute various activities to conduct and/or support successful benchmarking in an organization.

— ISO/IEC 29155-3 prescribes the guidance for reporting processes and contents of typical reports.

— ISO/IEC 29155-4 provides guidance for the activities to collect data of IT projects to be entered into and maintained in a benchmarking repository.

Further parts might follow.

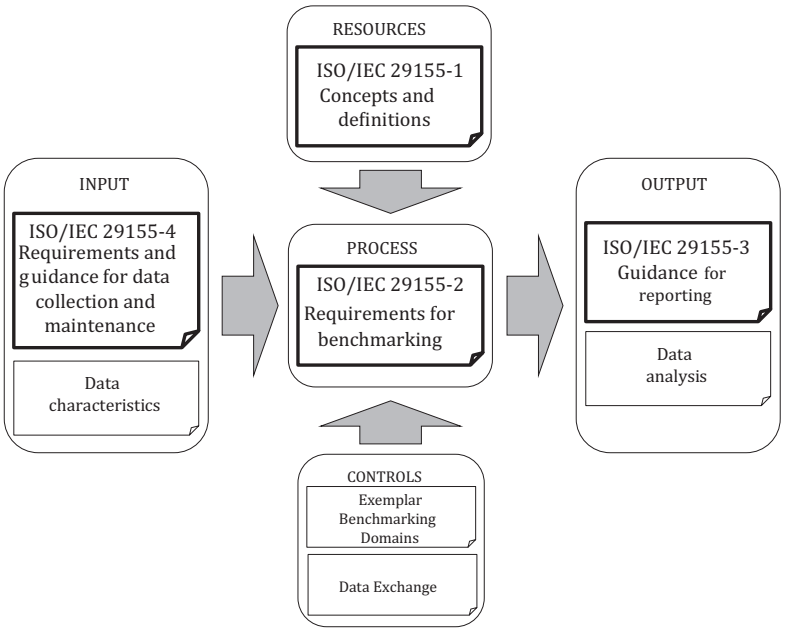


Figure 1 — IT project performance benchmarking standards overview

Systems and software engineering — Information technology project performance benchmarking framework —

Part 4: Guidance for data collection and maintenance

1 Scope

This document provides general requirements and guidance for collecting and maintaining data of information technology (IT) projects and for delivering the benchmarking repository within benchmarking activities of "the IT project performance benchmarking framework" by prescribing the following:

a) requirements and guidance for data element definitions;

b) requirements and guidance for the data collection and maintenance processes within the benchmarking framework;

c) requirements and guidance for maintaining benchmarking repository product and issued benchmarks.

This document mainly focuses on three major activities, which are "maintain repository", "submit IT project data", and "measure IT project" activities.

This document is intended for use by stakeholder(s) of IT project performance benchmarking (e.g. benchmarking user, benchmark provider, benchmarking service provider, and IT project team).

NOTE The following are examples of how this document can be used:

— by a benchmark provider, to define data elements, collect and maintain IT project data, and provide benchmarking repository product or issued benchmarks;

— by a benchmarking analyst, to use benchmarking repository product and/or benchmarks for executing an instance of benchmarking;

— by a benchmarking service provider, to utilize benchmarking repository product and/or benchmarks for providing benchmarking services;

— by an IT service provider, to define data elements to be measured and/or to be submitted to repository owner.

It is out of the scope of this document to prescribe a particular set of data element definitions, formats or contents of the benchmarking repository.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 29155-1, ISO/IEC 29155-2, and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp/

3.1
attribute
property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means

[SOURCE: ISO/IEC 15939:2007, 2.2]

3.2
data element
smallest unit of data of IT project

Note 1 to entry: Data element is usually implemented to be a data cell in a benchmarking repository and/or an "IT project data".

Note 2 to entry: Data element is defined by a responsible person who plans data collection or benchmarking.

3.3
data record
defined group of related *data elements* (3.2), in which all the necessary data elements are included to represent *attributes* (3.1) of interest

Note 1 to entry: One data record usually corresponds to a specific IT project within IT project performance benchmarking framework.

3.4
IT project dataset
classified group of *data records* (3.3), into which collected data records are selected by pre-defined criteria

Note 1 to entry: Classification criteria for an IT project dataset might be based on the information needs of the owner and/or users of data.

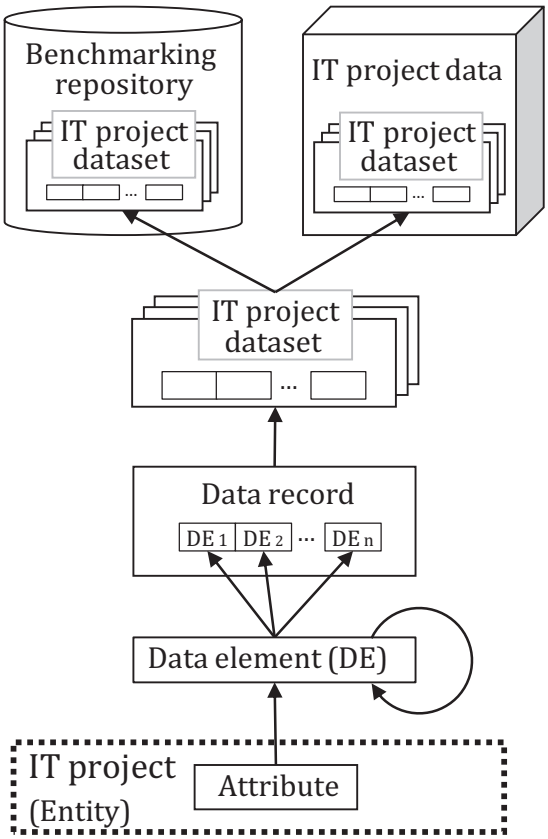
4 Abbreviated terms

IT information technology

5 Requirements and guidance for data definition

5.1 IT project data structure overview

In the framework of IT project performance benchmarking in the ISO/IEC 29155 series, there are two major data stores, which are "benchmarking repository" and "IT project data" (see Figure C.1). Figure 2 illustrates the overall data structure of these data stores.



NOTE To simplify this figure, only one instance of each construct is shown. In the actual implementation of the benchmarking framework, several instances of each construct might be defined.

Figure 2 — Data structure in the IT project performance benchmarking framework

An IT project is the entity of interest within a benchmarking framework that has various attributes to represent its property or characteristics.

An attribute is measured (or recorded) and stored into a data element. A data element is the smallest unit of accessing and manipulating data in the framework of the ISO/IEC 29155 series (see 5.3 for details). The owner of a "benchmarking repository" and/or "IT project data" defines data elements for measuring and maintaining data. Benchmarking analysts may also define data elements for their benchmarking purpose.

A data element is defined by prescribing its properties (see 5.6 for details).

NOTE 1 Examples of properties of a data element are: Identifier, name, data type, and measurement unit.

NOTE 2 In the case of data elements in the "IT project data", the purpose of data element definitions is not limited to the benchmarking activities. Data elements might also be defined to facilitate utilization of data in project management, quality assurance, financial management, and so on.

After defining data elements, one or more data records will be defined. A data record is a defined group of data elements, in which all the necessary data elements are included to represent attributes of interest of an IT project. One data record contains data of a specific IT project. Within the benchmarking framework, a data record is the typical object for maintaining and exchanging data of an IT project.

Data collection can be started after defining data elements and data record(s). In usual case, IT project is measured (or recorded) by IT project team, and data record(s) are submitted to the specific reception

desk or stored into a specified data storage. Collected data records may be collected and organized into several IT project datasets, which may be maintained independently.

NOTE 3 Classification criteria depend on the needs of the owner and/or users of data (see 5.5 for details).

Data in a benchmarking repository and/or an "IT project data" are retrieved by a benchmarking analyst for conducting an instance of benchmarking, or for issuing benchmark(s).

5.2 Attribute

An attribute is a property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means.

IT project is the major entity of interest in the ISO/IEC 29155 series.

NOTE In the ISO/IEC 29155 series, IT project attributes include, but are not limited to

— input for an IT project (e.g. requirements, resources, deliverables),

— outcomes of an IT project (which includes intermediate and temporary products together with final deliverables),

— information about an IT project itself (e.g. structures, processes, status, progress), and

— stakeholders (e.g. role, responsibility, expertise).

Usually, the first step in planning a specific instance of benchmarking information model is to select a set of IT project attributes that are relevant to the benchmarking needs. Selected attributes may be measured or recorded and maintained in the form of data elements to support various information needs in the future.

5.3 Data element

Within the IT project performance benchmarking framework, a data element is

a) the smallest unit of data of IT project,

b) usually implemented to be a data cell in a benchmarking repository and/or an "IT project data", and

c) defined by a responsible person who plans data collection or benchmarking.

NOTE 1 The responsible persons include, but are not limited to, benchmarking repository owner, IT project owner, benchmarking analyst, and benchmarking user.

For supporting benchmarking activities, benchmarking repository and "IT project data" usually contain explanatory information about the IT project (e.g. name and brief description of IT project, contact details of IT project team or data submitter) and/or its attributes (e.g. name of applied measurement method).

NOTE 2 See Annex A for more details of the nature of data elements.

5.4 Data record

Since tens or hundreds of data elements might be required to be measured and maintained within an IT project to meet multiple information needs of stakeholders, data elements and their order are usually organized.

A data record is a defined group of related data elements, in which all the necessary data elements are included to represent attributes of interest. One data record corresponds to a specific IT project.

The data record is the unit of measuring, maintaining, and exchanging data of an IT project in the benchmarking activities.

NOTE 1 Multiple data records might be defined from a set of defined data elements in an organization to satisfy individually different information needs. In such a case, a data element might be included in multiple data records.

NOTE 2 A data record is often defined in the form of an ordered set of data elements. Typical examples of such data record are a data table on spreadsheet and a table in a database,

EXAMPLE An IT project organization defined six data elements (project name, software size, total effort, total duration, number of test cases, and number of detected defects) to mandate IT project teams to measure and report. The quality assurance division of the organization defined a data record DR1 which consists of: project name, software size, number of test cases and number of detected defects. The project management office also defined a data record DR2 which consists of: project name, software size, total effort, and total duration. An IT project team measures six data elements and submits data in the form of required data records (DR1 for quality assurance division and DR2 for project management office).

5.5 IT project dataset

Data records in a data store (a benchmarking repository or an "IT project data"; see Figure C.1) may be bundled into group(s) to ease their usage and maintenance. Such groups are called IT project dataset within the IT project performance benchmarking framework.

An IT project dataset is a classified group of data records. The included data records might be selected by pre-defined criteria.

NOTE 1 Classification criteria for data record might be based on the information needs of the owner and/or users of data. Examples of dataset classification criteria are:

— data record definitions;

— project types(e.g. development, enhancement, and maintenance);

— business and/or application domain;

— organizational units;

— time period.

NOTE 2 A data record might be classified and stored into multiple IT project datasets.

NOTE 3 In implementing a physical data storage, multiple IT project datasets might be stored into a physical single data storage.

IT project datasets are stored and maintained in

— a benchmarking repository by the repository owner (e.g. benchmark provider or benchmarking service provider), and/or

— an "IT project data" by stakeholders of the IT project (e.g. IT project team, IT project owner, and benchmarking users).

5.6 Requirements for defining data elements

5.6.1 Overview of definitive properties

A data element is defined by prescribing a set of properties. A set of mandatory, recommended, and optional properties is summarized in Table 1. Additional properties may be prescribed to enhance definition of a data element.

Detail of properties in Table 1 are prescribed in 5.6.2 to 5.6.16.

Table 1 — Set of properties to define a data element

Property	Priority	Definition
<General properties>		
Identifier	[R]	Specify code or serial number of the data element.
Short name	[R]	Specify short name of the data element.
Name	[M]	Specify name of the data element.
Description	[M]	Specify overview information of the data element by describing meaning, purpose, and so on.
Priority	[O]	Specify priority of the data element (e.g. mandatory, recommended, optional).
Confidentiality	[O]	Specify condition of disclosure.
<Categorization properties>		
Reproducibility	[M]	Specify whether the value is reproducible or not.
Scalability	[M]	Specify whether the value is scalable or not.
<Measurement-related properties>		
Data type	[M]	Specify data type (e.g. integer, real number, string, enumeration) and its sub-properties.
Measurement unit	[Ms]	Specify measurement unit. [Applicable only for scalable data elements]
Precision	[Rs]	Specify required precision of value. [Applicable only for numeric data type]
Timing	[R]	Specify when the data element is produced.
Procedure	[M]	Specify how to produce the value of the data element, by describing logical sequence of operations (together with algorithms, if applicable), or by specifying a list of applicable existing procedures.
Source data elements	[Rr]	Specify the list of data elements which are necessary to produce value.
Dependency	[M]	Specify the list of data elements which are necessary to interpret value (e.g. qualifiers. See Annex A).

NOTE 1 Priority (upper case character): [M] Mandatory, [R]: Recommended, [O]: Optional.

NOTE 2 Priority (lower case character): [s]: Scalable data element, [r]: Reproducible data element, <none>: Applies to both.

5.6.2 Identifier

Data element should have a property called identifier. An identifier may be a code or a serial number.

If an identifier is specified, it shall be unique within IT project dataset.

NOTE Identifier is often referred for designing physical database or spreadsheet.

5.6.3 Short name

Data element should have a property called short name. If a short name is specified, it shall be unique within IT project dataset.

NOTE Users usually prefer short name rather than any technically generated identifier or long name, since it is easier to remember or refer.

5.6.4 Name

Data element shall have a property called name. A data element name shall be unique within an IT project dataset.

5.6.5 Description

Data element shall have a property called description.

Description should include, but is not limited to

— meaning of the data element,

— purpose of measurement (or recording), and

— relationship(s) to other data element(s), if applicable.

NOTE Description does not need to be exhaustive if the meaning, purpose or relationships are obvious from its name or other properties.

5.6.6 Priority

Data element may have a property called priority.

NOTE An example of priority rating is: "{mandatory, recommended, optional}".

5.6.7 Confidentiality

Data element may have a property called confidentiality. This property specifies the condition of disclosure.

Examples of confidentiality specification are

— the value of this data element shall be disclosed only to the benchmarking repository owner,

— the value of this data element shall be disclosed only to the quality assurance division staff, and

— the value of this data element shall be disclosed to only to the pre-authorized person.

5.6.8 Reproducibility

Data element shall have a property called reproducibility. This property designates whether the value of the data element is reproducible or not.

NOTE The value of a data element is reproducible if and only if it's source data elements are included in the data record.

5.6.9 Scalability

Data element shall have a property called scalability. This property designates whether the value of the data element has measurement scale or not.

NOTE Examples of descriptive data elements are: name of IT project, telephone number, textual explanation, photograph, and movies.

5.6.10 Data type

Data element shall have a property called data type. Table 2 shows examples of data types.

Table 2 — Typical data types and their exemplar sub-properties

Data type	Examples of sub-properties to be defined
Integer	Value range (e.g. in terms of acceptable minimum and maximum number), display format (e.g. numeral comma, plus and minus signs, zero suppression, value zero).
Real number	Value range (e.g. in terms of acceptable minimum and maximum number), recording precision (e.g. single, double, quadruple), display format (e.g. numeral comma, plus and minus signs, zero suppression, value zero, exponential part).
Currency	Value range (e.g. in terms of acceptable minimum and maximum number), display format (e.g. numeral comma, plus and minus signs, zero suppression, currency symbol).
Text string	Minimum and maximum length, character-code set (e.g. alphabetic, alpha-numeric, unlimited, exclude control-characters).
Date	Input format (order, separators, zero suppression, treatment of omitted elements), display format (order, separators, zero suppression, format of year).
Time	Input format requirement (order, separators, zero suppression, treatment of omitted elements), display format (order, separators, zero suppression, format of hour).
Enumeration (select from given list)	Complete list of alternatives, possibility of multiple selection, coding and displaying value and its format.
Code (input pre-defined code)	Access information for code table (e.g. URL, file name), input and display format (number of digits, separators, zero suppression).

NOTE In Table 2, Code is a data type which is systematically defined sequence of characters such as telephone or credit card number, ZIP code, country code, or user defined codes. It is not necessarily retrievable from pre-defined code table to input value.

5.6.11 Measurement unit

If the data element is scalable and it stores numeric value, the data element shall have a property called measurement unit. If the data element accepts one of multiple units, the list of acceptable units shall be presented. The selected and applied unit shall be recorded in another data element.

NOTE An example of list of acceptable units for [effort]: "{[person-hour], [person-day], [person-month], [person-year]}".

5.6.12 Precision

If the data element is scalable and it stores numeric value, data element should have a property called precision.

NOTE 1 An example of precision specification is: "Round off the value to two decimal places".

NOTE 2 The maximum precision of a reproducible data element usually depends on precision of its source data elements.

Input, display, and recording formats of a data element shall satisfy the precision specification.

NOTE 3 Displaying too many decimal digits might cause misunderstanding on actual precision of the data.

5.6.13 Timing

If user of collected data requires value of an attribute of an IT project in progress, data element shall have a property called timing to specify applicable time to produce data.

Timing is sometimes important since IT project attributes often change during the IT project lifecycle. If user requires a snapshot value of an attribute of an IT project in progress, this property shall be exactly specified.

EXAMPLE 1 A project manager wishes to conduct an instance of benchmarking at the end of system testing process to check [defect detection rate] for making decision to forward the program to next stage. To satisfy the benchmarking need, it is necessary to specify timing for [number of detected defects] and for [software size] to ensure that both measurement results represent the same stage of development.

Timing should be designated in terms of ISO/IEC standardized development lifecycle processes. If an organization-specific lifecycle is referred, it shall be mapped to ISO/IEC lifecycle process standard (see Example 2).

NOTE ISO/IEC standardized life cycle process is defined in ISO/IEC 12207 and ISO/IEC 15288.

EXAMPLE 2 Tables 3 and 4 show examples of mapping between an organization-specific lifecycle model and an ISO/IEC standard. Table 3 shows mapping of lifecycle phases of International Software Benchmarking Standards Group (ISBSG) to ISO/IEC 12207:2008, Table 4 shows mapping of lifecycle phases of Information-technology Promotion Agency (IPA) / Software quality Enhancement Center (SEC) to ISO/IEC 12207.

Table 3 — Example of mapping between an organization-specific lifecycle model and an ISO/IEC standard

No.	Phases of ISBSG	ISO/IEC 12207:2008
1	Specify	Requirements elicitation
		System requirements analysis
		System architecture design
		Software requirements analysis
2	Design	Software design
3	Build	Software construct (code & unit test)
		Software integration
4	Test	Software testing
		System integration
		System testing
5	Implement	Software installation
		User support

Table 4 — Example of mapping between an organization-specific lifecycle model and an ISO/IEC standard

No.	Phases of IPA/SEC	ISO/IEC 12207
1	Development planning	System development planning
2	Requirements definition	System requirements analysis
		Software requirements analysis
3	Basic design	System architecture design
		Software architecture design
4	Detailed design	Software detailed design
5	Development	Software coding and testing
6	System test	Software conformance test
		System conformance test
7	Acceptance test	Software installation support
		Software acquisition support
8	Follow-up (operation)	Operation process

5.6.14 Procedure

Data element shall have a property called procedure. A procedure is a logical sequence of operations (e.g. algorithms) to measure or record value of a data element. It may be specified by referring to existing procedures.

NOTE 1 An algorithm can be described in the form of mathematical formula.

NOTE 2 Examples of referable existing procedures are the following:

— apply one of the ISO/IEC standardized Functional Size Measurement Methods;

— encode picture by either {JPEG, JPEG2, GIF, TIFF, BMP}.

If a specified procedure produces multiple outcomes (e.g. both resulting values and intermediate values) the target value shall be clearly designated (see Example).

EXAMPLE Apply source-code checker XXX, and record the number of detected Level-2 defects.

If the data element can be produced using one of multiple procedures, the list of acceptable procedures shall be presented, and the selected and applied procedure shall be recorded in another data element.

5.6.15 Source data elements

Reproducible data element should have a property called source data elements. This property specifies the list of data elements which are necessary to produce the value of the data element.

NOTE An example of source data elements is: data elements [effort] and [size], which are referred for producing value of the data element [delivery_rate] (where [delivery_rate] = [effort]/[size]).

5.6.16 Dependency

A data element shall have a property called dependency. This property specifies the list of other data elements which are necessary to interpret its value (e.g. qualifiers; see Annex A).

5.7 Requirements for defining data records

The owners of data stores (benchmarking repository and “IT project data”) shall define data record(s) to be collected and maintained.

Data elements in a data record should be selected on the basis of the following:

— information needs (e.g. benchmarking) of stakeholders;

— feasibility and cost/benefit of data production;

— reproducibility and/or dependency.

NOTE 1 A process to determine a set of data elements to meet specific information needs is defined in ISO/IEC 15939.

NOTE 2 A reproducible data element becomes un-reproducible if any of its source data elements is not included in the data record. If reproducibility is lost it is necessary to revise the definition (e.g. description, procedure, source data elements) of the data element.

NOTE 3 If a data element is dependent to other data element(s), it is necessary to include those data element(s) together in the data record.

NOTE 4 Examples of data record definition are available in References [6] and [9].

If an organization intends to

— utilize external benchmarking repositories and/or external benchmarks, or

— submit data of IT project(s) to external organization(s),

compatibility should be taken into account to select and define data elements and data records.

NOTE 5 To minimize data collection costs, data elements in a data record might be prioritized into three categories as follows:

— [M] Mandatory, which is mandated to be collected and reported by all IT projects;

— [R] Recommended, which is not mandated but is strongly recommended to be collected and reported;

— [O] Optional, which collection and reporting is decided case by case.

5.8 Requirements for defining IT project datasets

The owners of data stores (benchmarking repository and "IT project data") should define IT project dataset(s) to enable high quality of management of unpredictable amount of data records in terms of

a) maintainability for owners of data, and

b) accessibility and availability for users of data.

NOTE 1 Examples of classifying data records into multiple IT project datasets for the owner of data are as follows:

— classify data records into three groups: before, during, and after the data validation;

— classify data records into two groups: completed IT projects and IT project in progress.

NOTE 2 Examples of classifying data records into multiple IT project datasets for the user of the data are as follows:

— classify data records by organizational units;

— classify data records by project score (e.g. succeeded, challenged, and canceled).

6 Functions for data collection

6.1 Overview

In the benchmarking framework of the ISO/IEC 29155 series, data collection is mainly performed in three activities:

— "measure IT project";

— "submit data";

— "maintain repository".

NOTE Activities of the benchmarking framework are defined in ISO/IEC 29155-1 (see Figure C.1), and tasks of each activity are given in ISO/IEC 29155-2.

6.2 Data collection functions of "Measure IT project"

In the benchmarking framework of the ISO/IEC 29155 series, the purpose of IT project measurement is not limited to use of data for the IT project performance benchmarking. "Measure IT project" activity is planned, and performed by several stakeholders of IT project, such as the following:

— by an IT project team or owner to know and manage progress of their project;

— by an IT project owner to know and manage performance and progress of IT project(s);

— by a benchmarking user to prepare data to conduct benchmarking;

— by an IT project related organization to submit data to the benchmarking repository owner.

Data collection functions in "Measure IT project" activity usually include, but are not limited to the following:

a) definition of the data elements, data record(s), and IT project dataset(s);

b) preparation of data collection tools (e.g. questionnaire, printed data entry form, computerized online/offline data entry form), data validation tools, and storage tools;

c) recruitment, reception, validation, and registration of data to "IT project data".

Since this activity is usually performed within an organization or within the closely related organizations, the organizational culture (e.g. commonly used data elements, data maintenance procedures, and measurement methods) is usually similar.

In such situation, if usage of IT project data is limited to be used within the organization(s), properties of data element are not necessarily so exhaustively prescribed. The pre-defined and commonly used rules (e.g. measurement unit, precision, timing, and procedures) may be referred to define data elements. Existing tools and procedures may also be applied to implement and support the "Measure IT project" activity. However, exact definition should be performed if measured IT project data are intended to be submitted to or used in external organization(s).

6.3 Data collection functions of "Submit data"

In the benchmarking framework of the ISO/IEC 29155 series, the purpose of IT project data submission is providing data of IT project(s) to be registered and utilized for IT project performance benchmarking.

Data collection functions in "Submit data" activity usually include, but are not limited to the following:

a) selection of IT project(s) to submit data;

b) preparation of data record(s) to be submitted;

c) submission of IT project data in the form of data record(s) or IT project dataset.

Such tasks as converting, normalizing, or re-measuring might be necessary before submission, if the definition of data elements or data record of required data is not precisely equivalent.

6.4 Data collection functions of "Maintain repository"

Functions in "Maintain repository" activity by the benchmark provider usually include, but are not limited to, the following:

a) definition of the data elements, data record(s), and IT project dataset(s);

b) preparation of data collection tools (e.g. questionnaire, printed data entry form, computerized online/offline data entry form), data validation tools, and storage tools;

c) recruitment, reception, validation, and registration of data of IT project(s) into a repository.

If IT project data are collected from culturally different organizations, definition of data elements should be as exhaustive as possible to avoid misunderstanding and/or unexpected measurement by data submitters. If several different definitions of a data element are popularly used across organizations (e.g. an effort is measured by person-months, person-days, and person-hours), convertibility should be carefully considered to merge and integrate data of IT projects from multiple organizations.

NOTE Sometimes, the benchmarking repository is referred as a "Master repository", which is the storage owned by the repository owner, managed by the repository administrator to store and maintain the collected data of IT projects. "Benchmarking repository product" is a subset of Master repository to be delivered as a product to be used.

7 Requirements and guidance for data collection

7.1 Requirements and guidance for data source

It is known that the performance of IT projects varies among business domains, software domains, scale of projects, development type (e.g. new development, enhancement, maintenance), success of projects (e.g. succeeded, challenged, canceled), and so on. The repository owner should carefully consider the amount and the quality of data to be collected to satisfy needs of target users of the benchmarking repository product and benchmarks.

Often, the benchmarking repository owners desire to serve as large audience as possible, which makes them willing to collect

— as many data records as possible,

— as many data elements as possible, and

— as complete and high quality data records as possible.

a) Volume, quality, and completeness of data often depends on the way to collect data. The ways of data collection can be categorized to three types: [selected] collect data from selected specific data source (e.g. organization, IT project team) by individually recruiting submissions,

b) [public] collect data from general public by recruiting submissions through public media,

c) [mandate] mandate data submissions to specified IT project team(s) or organization(s).

Table 5 — Expected influence of data collection type

Data collection type	Typical advantages	Typical disadvantages
[selected]	Can expect high quality data. Easy to control variety of data.	Amount of data limited.
[public]	Can expect large amount of data. Can expect wide variety in terms of completeness and quality of data.	Often occurs sampling-bias (e.g. only successful projects are submitted) Difficult to ensure data quality (e.g. validation is not easy across large audience)
[mandate]	Can expect large amount of data. Can expect full coverage of target audience.	Might include low quality data (e.g. because of low motivation to collect data).
NOTE Sampling-bias is systematic error due to e.g. a non-random sample of a population, causing some members of the population to be less likely to be included than others.		

The balance between advantages and disadvantages (see Table 5) should be considered when determining the way to collect data.

7.2 Guidance for timing

The timing of data collection can be

a) recruit and receive data in specific interval (e.g. once per year), and

b) always accept data through year.

The efforts of repository administrator (to validate and register) will be heavily unbalanced within the year in the former case. Therefore, it will be necessary in the former case to allocate large human resources within short period. In such a case, process and validation rules should be defined in detail to keep the quality of the administrator's work.

7.3 Guidance for data validation

The data collection functions (e.g. tasks of “submit data” and “maintain repository” activities) include several risks that might render collected data invalid, such as the following:

a) data elements are re-entered into a questionnaire or a form by submitter, and mistakes (e.g. input wrong value, enter into wrong field) might occur in this process;

b) submitter converts his/her data to fit to required unit of measurement, and mistakes might occur (e.g. calculation or precision errors);

c) submitter cannot provide a data element because of lack of information, but the data element might mistakenly be treated as “0” (zero) or “” (string of length zero) instead of “N/A” (not available);

d) wrong source data (e.g. incorrectly labelled data element or incorrect data population);

e) upstream data errors (e.g. incorrect data in original source);

f) incorrect data element definition (e.g. wrong dependency, wrong algorithms).

Three first risks on the list are examples of human errors, and the others of non-human.

Data validation is the best and most commonly used way to detect such errors. The repository owner should prepare a set of validation rules to ensure quality of the validation process and the resulting data.

Typical set of validation rules include, but is not limited to the following:

— check the data type to confirm entered value is represented by correct type;

— check the consistency between data elements (e.g. start_date ≤ end_date, total_value=∑detail_value);

— check the value range, whether the value is within reasonable range (e.g. size ≥ 0, percentage value is within [0 to 100]),

— check the value range by comparing the value and the average of existing data (e.g. typical productivity value of the existing IT projects is within the range of [9,10 to 23,4]).

NOTE 1 Examples of data validation rules are available in References [7] and [10].

To validate data efficiently and effectively, the repository owner should request the submitter to notify contact point for inquiry from the repository administrator.

As a result of validation, quality of the data should be evaluated and recorded.

NOTE 2 Examples of data quality indicator are as follows.

— A = The data submitted was assessed as being sound with nothing being identified that might affect its integrity.

— B = The submission appears fundamentally sound but there are some factors which could affect the integrity of the submitted data.

— C = Due to significant data not being provided, it was not possible to assess the integrity of the submitted data.

— D = Due to one factor or a combination of factors, little credibility should be given to the submitted data.

7.4 Requirements and guidance for privacy protection

It is important, and in many cases indispensable, to protect privacy of data submitters and all other stakeholders who require privacy protection.

NOTE 1 Target of privacy protection need to be expanded to the owners of systems and/or software when an IT project is developing customer-specific software or system and the project team has signed a non-disclosure agreement.

Privacy protection policy shall be documented and provided to data submitter. It usually includes, but is not limited to the following:

a) management processes of submitted data (e.g. submitted data are stored to enough secured computer, and only the authorized repository administrator and analyst can access);

b) usage restrictions of submitted data (e.g. data will be used only to publish the specific benchmark "XXX", submitted original data will never disclosed);

c) management processes of privacy related information (e.g. such information will be stored and managed with a special care, appreciating anonymity and security requirements, and accessed only by the authorized personnel).

NOTE 2 Examples of privacy related information are: submitter's name and contact information, project name, system name, software name, reference to the specific names such as development support tools.

NOTE 3 Examples of privacy protection policy documents are available in Reference [8].

8 Functions for maintenance and support

In the benchmarking framework of the ISO/IEC 29155 series, maintenance and support of benchmarking repository and products are mainly performed in two activities:

— "maintain repository";

— "issue benchmarks".

Maintenance and support functions in these activities include, but are not limited to, the following:

— maintenance of data in benchmarking repository;

— maintenance of benchmarking repository itself;

— delivery of benchmarking repository products;

— delivery of issued benchmarks;

— support for benchmarking products;

— support for benchmarking experience base.

NOTE Activities of the benchmarking framework are defined in ISO/IEC 29155-1 (see Figure C.1) and tasks of each activity are introduced in ISO/IEC 29155-2.

9 Requirements and guidance for maintenance and support

9.1 Requirements and guidance for data maintenance

Since data records of an IT project in the benchmarking repository are usually stored and utilized within a long period, data maintenance might be required in some occasions.

Data maintenance can be classified into two categories according to its degree of details: data element level maintenance and data record level maintenance.

Typical reasons for data maintenance at data element level are

a) an error of data value is detected,

b) the submitter requests to add and/or collect data value(s),

c) a change of data element and/or data record definition reflects to existing data in the repository, and

d) an organization-wide change (e.g. reorganizing a company, change of software lifecycle process) reflects to existing data in the repository.

NOTE An example of impact of organization-wide change is the renewal of organizational codes. It is necessary to update the value of the data element [Organizational_Code] of existing data record in the benchmarking repository to ensure accessibility of existing data records together with newly coming data records in the future.

In these cases, data changes for the benchmarking repository (i.e. master repository) shall be correctly reflected to every product (benchmarking repository products, benchmarks). If impacts of the change are serious, the notice should be distributed to the users of benchmarking repository products and benchmarks.

Typical examples of data maintenance at data record level are

— expired old data (e.g. over ten years from registration) are removed or managed separately,

— a huge dataset is divided into multiple datasets,

— remove specific data record(s) from repository, and

— a new data record is defined and initial dataset is created from existing data records.

9.2 Requirements and guidance for repository maintenance

A benchmarking repository shall be maintained for same reasons as any other typical database.

Typical exemplar operations for repository maintenance are

— backup and restore,

— user management and access control (authentication and authorization), and

— re-organization and re-configuration.

9.3 Requirements and guidance for delivery of benchmarking repository product

The benchmark provider delivers benchmarking repository products.

A benchmarking repository product is usually a subset of the benchmarking repository.

NOTE 1 The term “subset” here means that a product does not necessarily include all IT projects and/or all data elements from the “master” benchmarking repository.

In the production process,

a) all the privacy-related information shall be removed, and

b) explanatory report(s) shall be prepared.

NOTE 2 Requirements and guidance for the explanatory report is defined in ISO/IEC 29155-3.

The benchmarking repository product shall be delivered as a suite, which usually includes, but is not limited to the following:

— data product, which includes IT project dataset;

— explanatory report(s), which describe(s) data element definitions, user guide, contact information, and so on.

9.4 Requirements and guidance for delivery of issued benchmarks

The benchmark provider (e.g. the repository owner) may deliver benchmarks as their products.

The product may be

a) one or more benchmarks (e.g. values) which are analysed and derived from the benchmarking repository,

b) IT project dataset(s), which are subset(s) of the benchmarking repository.

In the production process of a benchmark product,

c) IT project dataset(s) shall be subset(s) of the benchmarking repository.

d) core report(s) shall be prepared to describe the issued benchmark(s) and an analysis process shall be used to derive benchmark(s), and

e) explanatory report(s) shall be prepared if core report(s) do not include the necessary information.

NOTE Requirements and guidance for a core report and an explanatory report are defined in ISO/IEC 29155-3.

9.5 Requirements and guidance for support for benchmarking products

The use of benchmarking products (e.g. benchmarking repository products, issued benchmarks) starts after their delivery, and usually continues at least to the delivery of the next version. Therefore, it is important to support the use of products also after the delivery.

After the delivery of the products, the support for its users should include the following functions:

a) responding to inquiries about the products;

b) correcting defects in the products (e.g. by replacing the products, by releasing errata or amendments);

c) receiving feedback from the users to improve processes of the activity.

9.6 Requirements and guidance for support for benchmarking experience base

To support the "improving" task group as outlined in ISO/IEC 29155-2, evaluation of the information products (e.g. a benchmark and a benchmarking report) and the benchmarking activities should be stored together with the lessons learned into the benchmarking experience base.

NOTE ISO/IEC 15939 defines "information product" as "one or more indicators and their associated interpretations that address an information need". For example, information product could include templates, graphs, statistical algorithms, and interpretation guidelines.

Annex A
(informative)

The benchmarking information model

A.1 Overview

The benchmarking information model is a structure that links benchmarking needs to the relevant IT projects and attributes of interest. The benchmarking information model describes how

— the relevant IT project attributes are measured (or recorded) and maintained in benchmarking repository and IT project datasets, and

— the archived data are processed to produce benchmarking outcomes (e.g. benchmarking reports and benchmarks) that provide a basis for utilizing benchmarking results.

The overall structure of the model is illustrated in Figure A.1.

NOTE 1 The benchmarking information model is built on the structure of the "measurement information model" which is defined in ISO/IEC 15939 to show relationships between information needs and target entities. The benchmarking needs correspond to the information needs, and the engineering and business objects of an IT project corresponds to the target entity in the benchmarking information model.

NOTE 2 The model in Figure A.1 is far more complex than that in ISO/IEC 15939:2007, Annex A. It is simply because the benchmarking repository or "IT project data" usually contains tens or hundreds of data elements to be able to meet various expected benchmarking needs, where information model in ISO/IEC 15939 treat only one or a few indicators at once to satisfy only one information need.

The selection or definition of appropriate data to address a benchmarking need begins with an idea of which measurable (or recordable) attributes of an IT project are related to benchmarking needs and how they are related.

The person planning data collection (e.g. benchmarking repository owner, IT project owner, or benchmarking analyst) defines components of the benchmarking information model that link IT project attributes to the specified benchmarking outcomes to satisfy the benchmarking needs. The benchmarking information model helps to determine what data and procedures the benchmarking stakeholders need to specify before starting benchmarking activities.

The benchmarking information model consists of two component groups:

— "data definition" group;

— "instance of benchmarking" group.

The "data definition" group consists of various data components (e.g. data element, data record, IT project dataset) and procedure components to manipulate them.

NOTE 3 Definition of the benchmarking repository is usually stable to keep continuity of data collection and services. In the other hand, definitions of "IT project data" are usually more dynamic because every IT project team might have different and dynamic information needs.

The "instances of benchmarking" group might be repeated frequently in "conduct benchmarking" activity. For this component group, the model in Figure A.1 shows very rough relationships from input data (benchmarking repository and "IT project data") to benchmarking outcomes. It is because procedures in this component group are very flexibly revised in every instance of benchmarking to meet ad hoc benchmarking needs.

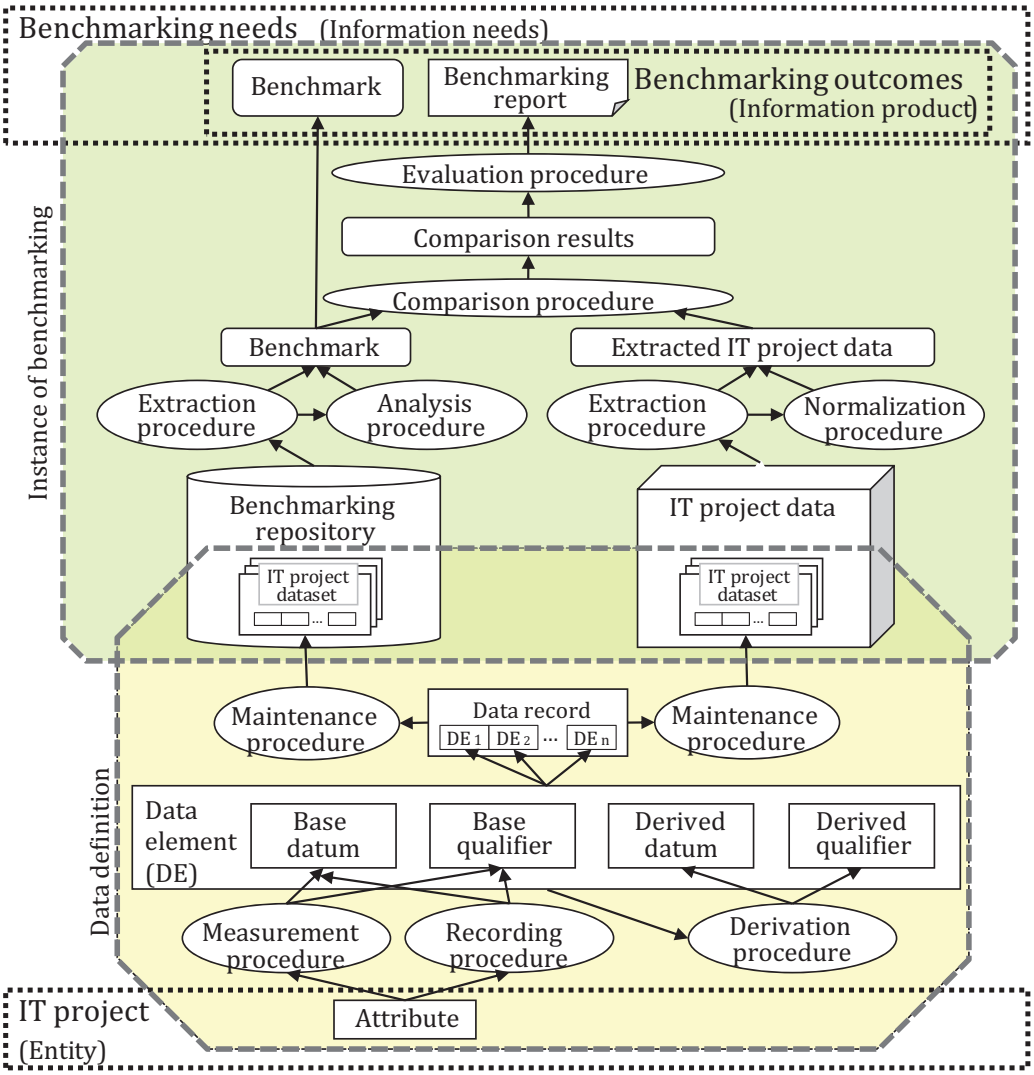


Figure A.1 — Benchmarking information model

A.2 Components of the model

A.2.1 IT project

In the benchmarking information model in the ISO/IEC 29155 series, an IT project is the object of benchmarking, and it is the major entity of interest.

NOTE According to the definition in ISO/IEC 15939, an entity is "an object that is to be characterized by measuring its attributes".

A.2.2 Attribute

An attribute is property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means.

NOTE 1 This definition is adopted from ISO/IEC 15939:2007, 2.2 (attribute).

IT project is the major entity of interest in the benchmarking information model, and it has various attributes to represent its property or characteristics.

Wide-ranging analyses might be required to address benchmarking information needs, and various kinds of multiple engineering and/or business attributes might be necessary to be measured or recorded. These attributes include, but are not limited to the following:

— input for IT project (e.g. requirements, resources, deliverables);

— outcomes of IT project (which includes intermediate and temporary products together with final deliverables);

— IT project itself (e.g. structures, processes, schedule, status, progress);

— stakeholders (e.g. role, responsibility, expertise).

Usually, the first step in planning a specific instance of the benchmarking information model is to select a set of IT project attributes that are most relevant to the benchmarking needs. A selected attribute may be measured or recorded in multiple ways and stored into multiple data elements to support different benchmarking needs in the future.

A.2.3 Data element

Within IT project performance benchmarking framework, a data element is

a) the smallest unit of data of IT project,

b) usually implemented to be a data cell in the benchmarking repository and/or the "IT project data", and

c) defined by a stakeholder who plans data collection or benchmarking,

NOTE 1 The stakeholders include, but are not limited to, benchmarking repository owner, IT project owner, benchmarking analyst, and benchmarking user.

From the difference of information source, a data element is classified into two groups:

— base data element: which is measured or recorded by applying a measurement procedure or a recording procedure to the attribute(s) of IT project;

— derived data element: which is produced by applying a derivation procedure to any one or more data elements.

NOTE 2 Derived data elements need not be stored into persistent storage since they are reproducible at any time by applying their derivation procedures.

For supporting benchmarking activities, benchmarking repository and "IT project data" usually contain explanatory information about the IT project (e.g. name and brief description of IT project, contact details of IT project team or data submitter) and/or its attributes (e.g. name of applied measurement method). This kind of information is usually recorded in the form of text string which does not have measurement scale. Therefore, from the viewpoint of measurement scale, a data element is either

— scalable data element: which is a "measure" and has measurement scale, or

— descriptive data element: which does not have measurement scale (e.g. name, telephone number).

NOTE 3 Examples of type of measurement scale are the following:

— **Nominal**: the values are categorical (e.g. classification of defects by their type);

— **Ordinal**: the values have rank orders. (e.g. the assignment of defects to a severity level);

— **Interval**: the values have equal distances corresponding to equal quantities of the attribute but the zero-point has no specific meanings (e.g. calendar date, temperature with the Celsius scale);

— **Ratio**: the measurement values have equal distances corresponding to equal quantities of the attribute where the value of zero corresponds to none of the attribute (e.g. age, effort, software size).

NOTE 4 Examples of descriptive data elements are [project name], [project description], and [name of the project manager].

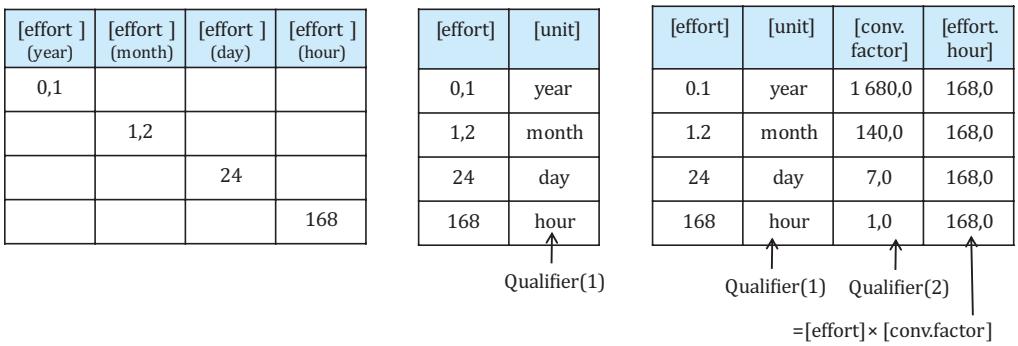
NOTE 5 A descriptive data element is not necessarily be free-format. It could have some recording rules (e.g. length of name is limited to less than 20 characters, only alpha-numeric characters could be used, telephone number need to start from country number with prefix "+").

A data element may be defined to be able to hold value by one of different scales (e.g. effort can be measured by one of person-hour, person-day, person-month). In such a case, it is necessary to define additional data element together to record supplementary information to interpret value of such data element. From this point of view, a data element can be

a) **datum**: which holds information of attribute(s) of IT project in the form of quantitative (i.e. scalable) or qualitative (i.e. descriptive) value, or

b) **qualifier**: which holds supplementary information to interpret value of one or more data elements (e.g. measurement method, measurement precision, measurement unit, measurement timing, scale conversion factor). It also may be in the form of quantitative or qualitative value.

EXAMPLE Four values [effort (unit: person-year)], [effort (unit: person-month)], [effort (unit: person-day)] and [effort (unit: person-hour)] are logically different, since they are measured with different scales. If these are stored in different data cells, resulting data table will have four rows [see Figure A.2 a)]. This table is practically very inefficient, since value is stored in only one of the four cells and the table becomes very sparse. Introducing a qualifier [unit (of effort)] solves the issue to refine table to be efficient [see Figure A.2 b)]. In addition, introducing another qualifier [conv. factor (to person-hour)] will enable user to compare values with different scale by adding a derived data element [effort (by-person) hour] [see Figure A.2 c)].



a) Sparse table **b) Refined table (1)** **c) Refined table (2)**

Figure A.2 — Exemplar usage of qualifier elements

A.2.4 Measurement procedure

A measurement procedure measures attribute(s) of an IT project to produce scalable base datum. A measurement procedure may also produce base qualifier(s) to record supplementary information to interpret value of specific datum. Complex measurement procedure may also produce descriptive datum. Possible input(s) and output(s) of a measurement procedure are illustrated in Figure A.3.

A measurement procedure is a logical sequence of operations used in quantifying and/or qualifying specific attribute(s). The operations usually include, but are not limited to, tasks such as counting occurrences or observing the passage of time.

The same measurement procedure may be applied to multiple attributes to measure multiple base data elements.

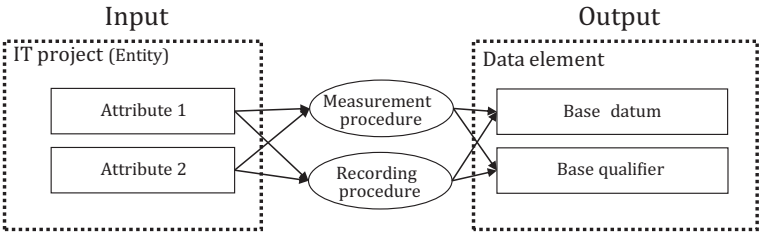


Figure A.3 — Input and output of measurement procedure and recording procedure

A.2.5 Recording procedure

A recording procedure records descriptive base datum and base qualifier(s).

A recording procedure is a logical sequence of operations, used in recording properties of attribute(s) of an IT project into descriptive base datum. The operations may include activities, but are not limited to, typing texts, taking photographs, drawing pictures, recording movies or sounds. The same recording procedure may be applied to multiple attributes for recording multiple descriptive base data elements.

A recording procedure may also produce base qualifier(s) to record supplementary information to interpret value of specific data element(s).

A.2.6 Derivation procedure

A derivation procedure is a logical sequence of operations to produce derived element(s).

Any data element(s) may be input for a derivation procedure. The output(s) of derivation procedure are always derived data element(s) (see Figure A.4).

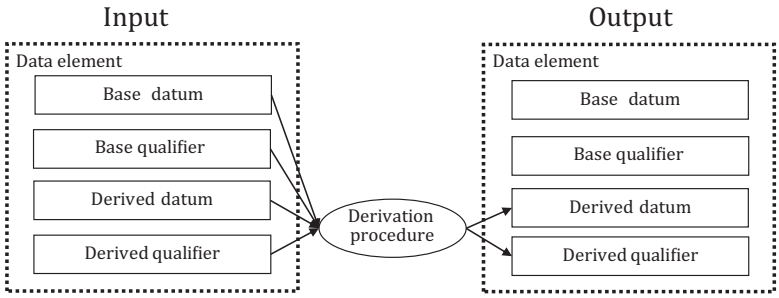


Figure A.4 — Input and Output of Derivation procedure

A.2.7 Data record

A data record is a defined set of related data elements, in which all the necessary data elements are included to represent attributes of interest of an IT project. One data record corresponds to a specific IT project. The data record is the unit of measuring, maintaining, and exchanging data of an IT project in the benchmarking activities.

NOTE 1 Multiple different data records might be defined from a set of defined data elements in an organization to satisfy individually different information needs. In such a case, a data element might be included in multiple data records.

NOTE 2 A data record is often defined in the form of an ordered-set of data elements. Typical examples of such data record are a data table on spreadsheet software and a table in a database.

A.2.8 IT project dataset

Since the benchmarking repository and IT project data usually store data of large amount of IT projects, data records may be bundled into group(s) to ease their usage and maintenance.

An IT project dataset is a classified group of data records, into which collected data records are selected by pre-defined criteria.

NOTE 1 Classification criteria depend on the information needs of the owner and/or user of data. Examples of classification criteria are categorizing data of IT projects by

— data record definitions,

— project types(e.g. development, enhancement, and maintenance),

— business and/or application domain,

— organizational units or IT project teams, and

— selecting IT projects within specific time period.

NOTE 2 An data record might be classified and stored into multiple IT project datasets.

NOTE 3 In implementing physical data storage, multiple IT project datasets might be stored into the physically single data storage by attaching a qualifier to distinguish which dataset(s) a data record belongs to.

IT project datasets are stored and maintained in

— the benchmarking repository by the repository owner (e.g. benchmark provider or benchmarking service provider), or

— the “IT project data” by stakeholders of IT project (e.g. IT project team, IT project owner, and benchmarking users).

A.2.9 Maintenance procedure

A maintenance procedure registers and maintains data records in IT project datasets. A tasks of maintenance procedure include, but are not limited to

— adding, updating, deleting data records in IT project dataset(s),

— backup and/or restore data records in IT project dataset(s),

— creating and/or modifying data element or data record definitions,

— optimizing physical data storage, and

— moving or copying data records to other IT project dataset.

A.2.10 Extraction procedure

An extraction procedure selects a subset of IT project dataset from the benchmarking repository or “IT project data”. In the extraction process, a subset is created by

a) ignoring unnecessary data elements in data element groups, and/or

b) selecting only specific IT projects of interest that match the selection criteria.

Outcome of an extraction procedure might be a benchmark or extracted IT project dataset. In some cases, providing a benchmark requires specific analysis, and providing extracted data in IT project dataset requires normalization.

NOTE In the framework of the ISO/IEC 29155 series, both benchmark and extracted IT project data might be

— numeric value(s) (e.g. value of data element, statistic),

— IT project dataset(s) of selected IT projects,

— statistical graph(s) (e.g. scatter graph, histogram), and so on.

A.2.11 Analysis procedure

A analysis procedure is a logical sequence of operations to derive a benchmark.

NOTE Statistical analysis techniques are often applied to derive benchmarks.

A.2.12 Normalization procedure

A normalization procedure is a logical sequence of operations to make data comparable.

NOTE 1 Examples of reasons to normalize data include

— different data definitions,

— missing data values,

— different measurement units, and

— inconsistent data collection.

NOTE 2 Examples of normalization include

— converting the measurement unit and/or value of a data element,

— adding an un-featured derived data element to IT project dataset (if input data elements are available), and

— adding an base element and populating it through some logical or mathematical function (e.g. aggregating, dividing, interpolating).

A.2.13 Comparison procedure

A comparison procedure is a logical sequence of operations to compare the extracted IT project data against the benchmark.

Comparison is based on a set of assumptions and expected relationships between the benchmark and the extracted IT project dataset. The procedure provides a collection of observations and findings that are expected to be relevant to the given benchmarking needs. Typical comparison results are represented in the form of indicator(s), statistical chart(s), textual description(s) and so on.

A.2.14 Evaluation procedure

An evaluation procedure examines the applicability of the comparison results to satisfy the benchmarking needs in term of the pre-defined judgment criteria. An evaluation procedure might be performed by the benchmarking analyst and/or be automated.

NOTE In the ISO/IEC 29155 series, judgment method and criteria are required to be defined in the planning phase of “conduct benchmarking” activity (see ISO/IEC 29155-2).

If the comparison result is judged that it satisfies judgment criteria, the benchmarking analyst produces a benchmarking report and submits it to benchmarking user.

Annex B
(informative)

Classification of data elements and comparison with other related standards

B.1 Overview

This annex classifies data elements in the benchmarking information model in the ISO/IEC 29155 series into eight categories from the three aspects, which are target of measurement, production process, and level of measurement scale.

This annex also compares characteristics of each category's data element against measures in ISO/IEC 15939 or QMEs in ISO/IEC 25021. Table B.1 outlines the overall differences between three International Standards.

Table B.1 — Comparison of characteristics of data elements with related standards

	Data element in ISO/IEC 29155 series	Measure in ISO/IEC 15939	Quality Measure and QME in ISO/IEC 25021
Definition	**data element:** a) The smallest unit of IT project data. b) Usually implemented to be a data cell in the benchmarking repository and/or the IT project data.	**measure:** Variable to which a value is assigned as the result of measurement.	**quality measure:** Derived measure that is defined as a measurement function of two or more values of quality measure elements. **quality measure element (QME):** Measure defined in terms of a property and the measurement method for quantifying it, including optionally the transformation by a mathematical function.
Nature	Practical (physical).	Logical (theoretical).	Logical (theoretical).
Variation	Base datum Derived datum Base qualifier Derived qualifier	Base measure Derived measure Indicator	Quality measure Quality measure element
Descriptive data without scale	Included within scope as descriptive data element.	Out of scope.	Out of scope.
Qualifier	Included within scope as qualifier.	Necessity is mentioned.	Not mentioned.
Configuration	Instantiated as a data cell in benchmarking repository or IT project data to store value.	Defined as a logical variable but is not instantiated.	Defined as a logical variable but is not instantiated.
Data structure	IT project dataset consists of a set of data elements to represent an IT project.	No data structure are defined.	No data structure are defined.

Table B.1 *(continued)*

	Data element in ISO/IEC 29155 series	Measure in ISO/IEC 15939	Quality Measure and QME in ISO/IEC 25021
Data management	A set of data elements in IT project dataset	Not mentioned.	Not mentioned.
Data type	Data type and precision are explicitly defined.	Not explicitly mentioned.	
Data definition requirements	Prescribed in detail as the data element definition requirements.	Only exemplar definitions are shown in informative annex.	Prescribed in detail as the QME definition requirements.

B.2 Classification of data elements

B.2.1 General

Data elements in the ISO/IEC 29155 series can be classified by three aspects as follows:

— target of measurement (or recording);

— production process;

— level of measurement scale.

Since these three aspects can be applied to every data elements, there are eight (= 2 × 2 × 2) data element categories as is shown below (see Figure B.1).

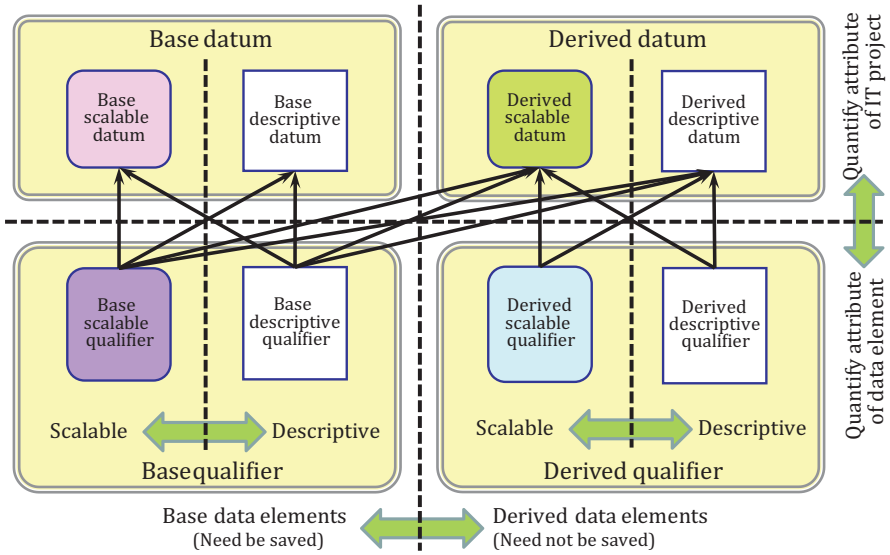


Figure B.1 — Classification of data elements in the ISO/IEC 29155 series

B.2.1.1 Aspect — Target of measurement (or recording)

Data elements are classified to

— **datum**: a data element which quantifies (or records) attribute of IT project, and

— **qualifier**: which quantifies (or records) supplementary information to interpret value of one or more data element(s).

NOTE 1 A qualifier is measured, recorded, or derived (by measurement method, recording method, or derivation procedure) together with other data elements.

NOTE 2 Examples of qualifier for the datum: [effort of acceptance test process] are

— [unit of effort] which is indispensable for archiving data if effort data are measured by multiple units within data provider, and

— [conversion factor (hours per month)] which is indispensable for data user to normalize effort data to desired scale or unit.

NOTE 3 A qualifier can qualify multiple data elements. An example of such case is that a qualifier: [unit of effort] qualifies data elements: [effort of XXXX process].

B.2.1.2 Aspect — Production process

Data elements in the ISO/IEC 29155 series can also be classified to

— **base data element**: a data element that is measured (or recorded) by applying measurement (or recording) method to IT project attribute(s), and

— **derived data element**: a data element that is defined in the form of derivation procedure.

This classification is applicable for both datum and qualifier.

NOTE 1 Derived data elements are re-producible at any time by applying derivation procedure, therefore they need not to be stored in persistent storage.

NOTE 2 The ISO/IEC 29155 series does not distinguish “indicator” with derived data element in the aspect. The term “indicator” does not appear in Figure A.2, since result of benchmarking usually cannot be explained by a single numeric value.

B.2.1.3 Aspect — Level of measurement scale

Data elements can be also classified (see Figure B.2) to

— **scalable data element**: a data element that has scale, and

— **descriptive data element**: a data element that does not have scale.

This classification is also applicable for both datum and qualifier.

NOTE 1 Descriptive data is indispensable information in benchmarking to record explanatory information of data (e.g. profile information of IT project, identification of IT project, or contacts of data submitter).

NOTE 2 The ISO/IEC 29155 series avoid to use term “measure” (both as verb and noun) to descriptive data elements. Descriptive data element is not a “measure” (noun) in the measurement theory, and it is “recorded” instead of “measured”.

NOTE 3 Descriptive data is out of scope of ISO/IEC 15939 and ISO/IEC 25021.

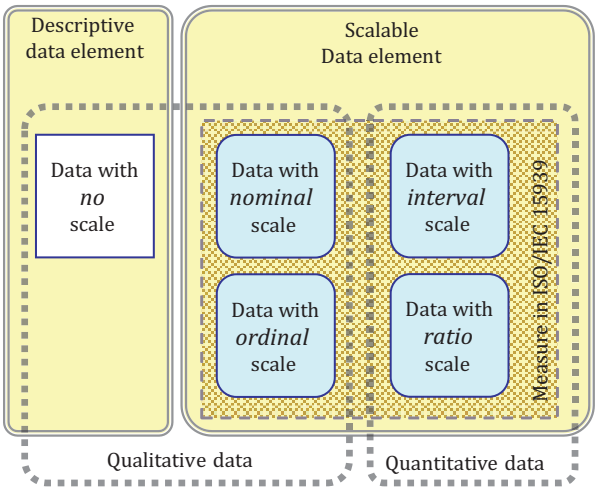


Figure B.2 — Level of measurement scale of data elements

B.2.2 Base scalable datum

A base scalable datum is the data element that

— **[target of measurement]**: quantifies IT project attribute(s),

— **[production process]**: is measured by applying a specific measurement method by human or automated means, and

— **[level of measurement scale]**: is scaled value.

NOTE A base data element of this type is classified to the "base measure" in ISO/IEC 15939:2007, 2.3.

B.2.3 Base descriptive datum

A base descriptive datum is the data element that

— **[target of measurement]**: records IT project attribute(s),

— **[production process]**: is recorded by applying a specific recording method by human or automated means, and

— **[level of measurement scale]**: is value with no scale.

NOTE 1 Examples of base descriptive datum are [project name], [project description], and [leader name].

NOTE 2 A base data element of this type is not the "base measure" in ISO/IEC 15939:2007, 2.3 since it does not have scale.

B.2.4 Base scalable qualifier

A base scalable qualifier is the data element that

— **[target of measurement]**: preserves supplementary information to interpret value of specific data element(s),

— **[production process]**: is quantified together with base data element(s) by applying a specific measurement method by human or automated means, and

— **[level of measurement scale]**: is scaled value.

NOTE 1 A base scalable qualifier initially qualifies only one or more base data elements. Qualification relationship can be inherited to derived data elements when context information of a qualified base data element is succeeded to derived elements.

NOTE 2 An example of base scalable qualifier is conversion factor [work hours per day] to convert unit of [effort] from [person-day] to [person-hour].

NOTE 3 A base qualifier of this type is classified to the "base measure" in ISO/IEC 15939:2007, 2.3.

B.2.5 Base descriptive qualifier

Base descriptive qualifier is the data element that

— **[target of measurement]**: preserves supplementary information to interpret value of specific data element(s),

— **[production process]**: is recorded together with base data element(s) by applying a specific recording method by human or automated means, and

— **[level of measurement scale]**: is value with no scale.

NOTE 1 A base descriptive qualifier initially qualifies only one or more base data elements. Qualification relationship can be inherited to derived data elements when context information of a qualified base data element is succeeded to derived elements.

NOTE 2 An example of base descriptive qualifier element is [name of applied counting tool] in free format. If the candidate tools are limited, this data element can be defined as choice from enumerated list (i.e. nominal scale).

NOTE 3 A base qualifier of this type is not the "base measure" in ISO/IEC 15939:2007, 2.3 since it does not have scale.

B.2.6 Derived scalable datum

A derived scalable datum is the data element that

— **[target of measurement]**: quantifies IT project attribute(s),

— **[production process]**: is derived from one or more data elements (of any kind) by applying derivation procedure (together with referring or succeeding context information from), and

— **[level of measurement scale]**: is scaled value.

NOTE 1 Transformation of one data element (of any kind) is deemed to be an new derived scalable datum in the ISO/IEC 29155 series since such conversion (e.g. change of measurement unit) gives significant value in benchmarking.

A derived scalable datum can be classified into three types from the definitions in ISO/IEC 15939 and ISO/IEC 25021 as follows:

— **[Type-1]** a derived scalable datum that is produced from two or more base scalable datum;

NOTE 2 This type is classified to a "derived measure" in both ISO/IEC 15939 and a "quality measure" in ISO/IEC 25021.

NOTE 3 An example of this type is [defect density rate] := [number of detected defect]/[software size].

— **[Type-2]** a derived scalable datum that is produced from only one base scalable datum;

NOTE 4 An example of this type is a transformation of unit of data element: [effort] from [person-hour] to [person-month].

NOTE 5 This type is not classified to a "derived measure" in ISO/IEC 15939. ISO/IEC 15939:2007, A.2.4 says that *"Simple transformations of base measures (for example, taking the square root of a base measure) do not add information, thus do not produce derived measures."*. In the other hand, it is deemed to be a new "quality measure" (derived measure) in ISO/IEC 25021:2012, definition 4.14.

— **[Type-3]** a derived scalable datum that is produced from only one or more base or derived descriptive data element(s).

NOTE 6 An example of this type is a derivation procedure to measure and store "number of characters" of [project name] (which contains text string). The value [project name] is a scaleless value, and it is measured at the first time by derivation procedure.

NOTE 7 This type is classified to a "base measure" in both ISO/IEC 15939 and ISO/IEC 25021, since it can be interpreted that the measurement against the IT project attribute(s) is performed by derivation procedure at the time.

B.2.7 Derived descriptive datum

A derived descriptive datum is the data element to store scaleless value that is produced by applying derivation procedure to any data element(s).

A derived descriptive data element is the data element that

— **[target of measurement]**: records IT project attribute(s),

— **[production process]**: is derived from one or more (base or descriptive) data elements by applying derivation procedure, and

— **[level of measurement scale]**: is value with no scale.

NOTE 1 Examples of this type are textual operations (e.g. concatenation, re-formatting, changing character-code) to base or derived descriptive data elements.

NOTE 2 A derived data element of this type is not the "measure" in ISO/IEC 15939 since its value does not have scale.

B.2.8 Derived scalable qualifier

A derived scalable qualifier is the data element that

— **[target of measurement]**: records supplementary information to interpret value of specific derived data element(s),

— **[production process]**: is produced together with any derived data element(s) from one or more data elements of any categories by applying derivation procedure, and

— **[level of measurement scale]**: is scaled value.

A derived scalable qualifier can be classified into two types in the definitions of ISO/IEC 15939 and ISO/IEC 25021 as follows.

— **[Type-1**] a derived scalable qualifier that is produced together with Type-1 of derived scalable datum by applying derivation procedure to preserve supplementary information to interpret value of specific derived data element(s).

NOTE 1 A derived qualifier of this type is classified to the "derived measure" in ISO/IEC 15939:2007, 2.3.

— **[Type-2**] a derived scalable qualifier that is produced together with Type-2 or Type-3 of derived scalable datum by applying derivation procedure to preserve supplementary information to interpret value of specific derived datum.

NOTE 2 A derived data element of this type is classified to the "base measure" in ISO/IEC 15939:2007, 2.3.

B.2.9 Derived descriptive qualifier

A derived descriptive qualifier is the data element that

— **[target of measurement]**: records supplementary information to interpret value of specific base or derived datum,

— **[production process]**: is produced together with any derived data element(s) from one or more data elements of any categories by applying derivation procedure, and

— **[level of measurement scale]**: is value with no scale.

NOTE 1 Examples of this type are textual operations (e.g. concatenation, re-formatting, changing character-code) to base or derived descriptive qualifier element(s).

NOTE 2 A base qualifier of this type is not the “base measure” in ISO/IEC 15939:2007 2.3 since it does not have scale.

Annex C
(informative)

Framework model of IT project performance benchmarking in ISO/IEC 29155-1

This annex outlines the overall framework of IT project performance benchmarking (see Figure C.1).

NOTE See ISO/IEC 29155-1 for more information about the benchmarking framework and its concepts and definitions.

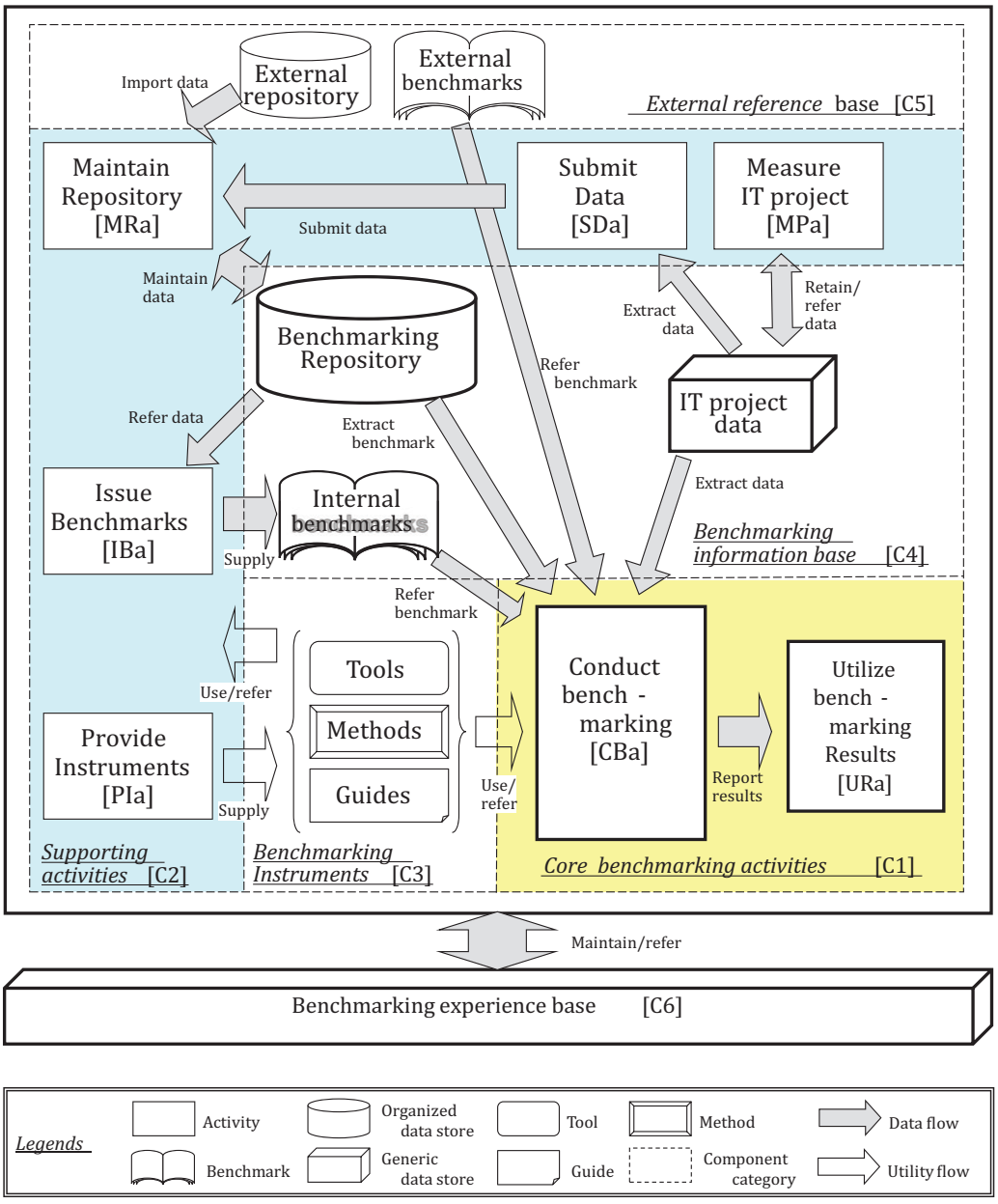


Figure C.1 — IT project performance benchmarking framework

The framework consists of the following six categories of activities or components:

— **[C1] "Core benchmarking activities"**, which collectively illustrates an instance of benchmarking;

— **[C2] "Supporting activities"**, which provides the benchmarking information base as well as instruments for use in an instance of benchmarking;

— **[C3] "Benchmarking instruments"**, used by stakeholders to conduct an instance of benchmarking;

— **[C4] "Benchmarking information base"**, which contains data for use during an instance of benchmarking;

— **[C5] "External reference base"**, which provides alternative or additional external data (e.g. repository and/or benchmarks) for an instance of benchmarking;

— **[C6] "Benchmarking experience base"**, which contains knowledge and lessons learned from present and/or past benchmarking experiences.

The first two categories (C1 and C2) introduce activities to conduct or support an instance of benchmarking. They consist of seven activities:

a) **"Conduct benchmarking"** activity, which initiates and executes an instance of benchmarking [CBa];

b) **"Utilize benchmarking results"** activity, which utilizes benchmarking results for various business purposes [URa];

c) **"Measure IT project"** activity, which measures IT project artifacts and maintains data [MPa];

d) **"Submit data"** activity, which selects and provides data of IT projects to be included into the benchmarking repository [SDa];

e) **"Maintain repository"** activity, which accepts, verifies, and stores IT project data into the benchmarking repository, and later manages the benchmarking repository [MRa];

f) **"Issue benchmarks"** activity, which analyses IT project data within the benchmarking repository and provides internal benchmarks [IBa];

g) **"Provide instruments"** activity, which develops and provides tools, methods, and guides to support every activity in the benchmarking framework [PIa].

For effective benchmarking, all seven of the above activities need to be properly executed.

The third category (C3) defines three types of benchmarking instruments:

— **Tools**, which provide means to support stakeholders;

— **Methods**, which provide procedures;

— **Guides**, which provide instructive information.

The remaining three categories (C4, C5, and C6) consist of various types of information bases:

— **Benchmarking repository**, which is a data store to maintain reliable data of IT projects used to produce a benchmark;

— **IT project data**, which is a data store to maintain data related to various IT projects;

— **Internal benchmarks,** which provide a group of commonly usable and authorized pre-determined benchmarks;

— **External repository**, which is a repository maintained by an external organization;

— **External benchmarks**, which are issued in some external organizations by analysing external repositories;

— **Benchmarking experience base** is an archive of benchmarking outcomes and lessons learned.

In the context of the ISO/IEC 29155 series, the concepts of "internal" and "external" are subjective views by stakeholders, and it is dependent on the judgment of the stakeholder. For example, if Organization X decides to adopt IT project data of Organization Y, then the data of Organization Y is considered to be a part of Organization X's internal data. Such adoption of out-of-organization data or benchmark often occurs within an enterprise group or between closely related business partners.

Bibliography

[1] ISO/IEC 12207:2008, *Systems and software engineering — Software life cycle processes*

[2] ISO/IEC 15288, *Systems and software engineering — System life cycle processes*

[3] ISO/IEC 15939:2007, *Systems and software engineering — Measurement process*

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

[5] ISO/IEC 25021:2012, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality measure elements*

[6] 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/

[7] 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/

[8] ISBSG. *Confidentiality and integrity of the data*, International Software Benchmarking Standards Group(ISBSG), http://www.isbsg.org/

[9] ISBSG. *Field Descriptions: Data Release 11*, International Software Benchmarking Standards Group(ISBSG), freely available at http://www.isbsg.org/

[10] ISBSG. *Data validation standard*, International Software Benchmarking Standards Group(ISBSG), freely available at http://www.isbsg.org/

[11] ISO/IEC 29155-1, *Systems and software engineering — Information technology project performance benchmarking framework — Part 1: Concepts and definitions*

[12] ISO/IEC 29155-2, *Systems and software engineering — Information technology project performance benchmarking framework — Part 2: Requirements for benchmarking*

ICS 35.080
Price based on 35 pages