INTERNATIONAL STANDARD

ISO/IEC 25066

First edition
2016-06-15

Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for Usability — Evaluation Report

Ingénierie des systèmes et du logiciel — Exigences de qualité et évaluation des systèmes et du logiciel (SQuaRE) — Format de l'industrie commune pour l'utilisation — Rapport d'évaluation

Reference number
ISO/IEC 25066:2016(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 document 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 and IEC 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 WTO principles in the Technical Barriers to Trade (TBT), see the following URL: Foreword — Supplementary information.

The committee responsible for this document is ISO/TC 159, *Ergonomics*, Subcommittee SC 4, *Ergonomics of human-system interaction* and Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 7, *Software and system engineering*.

Introduction

The human-centred design approach of ISO 9241-210 is well established and focuses specifically on making systems usable. Usability can be achieved by applying human-centred design throughout the lifecycle. In order to enable a human-centred approach to be adopted, it is important that all the relevant types of information related to usability (information items) are identified and communicated. The identification and communication of relevant types of information related to usability enables the design and testing of the usability of a system.

This International Standard provides a framework and consistent terminology for reporting the evaluation of an interactive system. It is intended to assist usability specialists and developers in documenting and communicating usability-related information as part of the system development lifecycle.

The Common Industry Format (CIF) for Usability family of International Standards is described in ISO/IEC TR 25060 and is part of the SQuaRE (Systems and software Quality Requirements and Evaluation) series of standards on systems and software product quality requirements and evaluation (ISO/IEC 25000[1)], ISO/IEC 25001, ISO/IEC 25021[2)], ISO/IEC 25023[3)], ISO/IEC 25040, ISO/IEC 25041 and ISO/IEC 25051).

The CIF family of standards uses definitions that are consistent with the ISO 9241 series of standards (Ergonomics of human-system interaction), as this is the terminology that is normally used for this subject matter. In some cases, these definitions differ from those in ISO/IEC 25000.

CIF standards are published or planned for the following information items:

— Common Industry Format (CIF) for usability test reports (ISO/IEC 25062);

NOTE ISO/IEC 25062 provides more detail for the content of a user observation report for performance measurement.

— Context of use description (ISO/IEC 25063);

— User needs report (ISO/IEC 25064);

— User requirements specification (ISO/IEC 25065);

— Evaluation reports (ISO/IEC 25066);

— User interaction specification (planned);

— User interface specification (planned);

— Field data report (planned).

The CIF standards are part of the “Extension Division” of the ISO/IEC 25000 SQuaRE series of International Standards. Table 1 presents an overview of the structure and the contents of the SQuaRE series of International Standards.

1) Withdrawn.

2) Withdrawn.

3) Under development.

Table 1 — Organization of SQuaRE series of International Standards

<table>
<tr><th colspan="3">SQuaRE Architecture and Sub-projects</th></tr>
<tr><td rowspan="3">ISO/IEC 2503n: Quality Requirement Division</td><td>ISO/IEC 2501n: Quality Model Division</td><td rowspan="3">ISO/IEC 2504n: Quality Evaluation Division</td></tr>
<tr><td>ISO/IEC 2500n: Quality Management Division</td></tr>
<tr><td>ISO/IEC 2502n: Quality Measurement Division</td></tr>
<tr><td colspan="3">ISO/IEC 25050 – 25099 SQuaRE Extension Division</td></tr>
<tr><td colspan="2">ISO/IEC 25051: Requirements for quality of Ready to Use Software Product (RUSP) and instructions for testing</td><td>ISO/IEC 2506n Common Industry Format Division</td></tr>
</table>

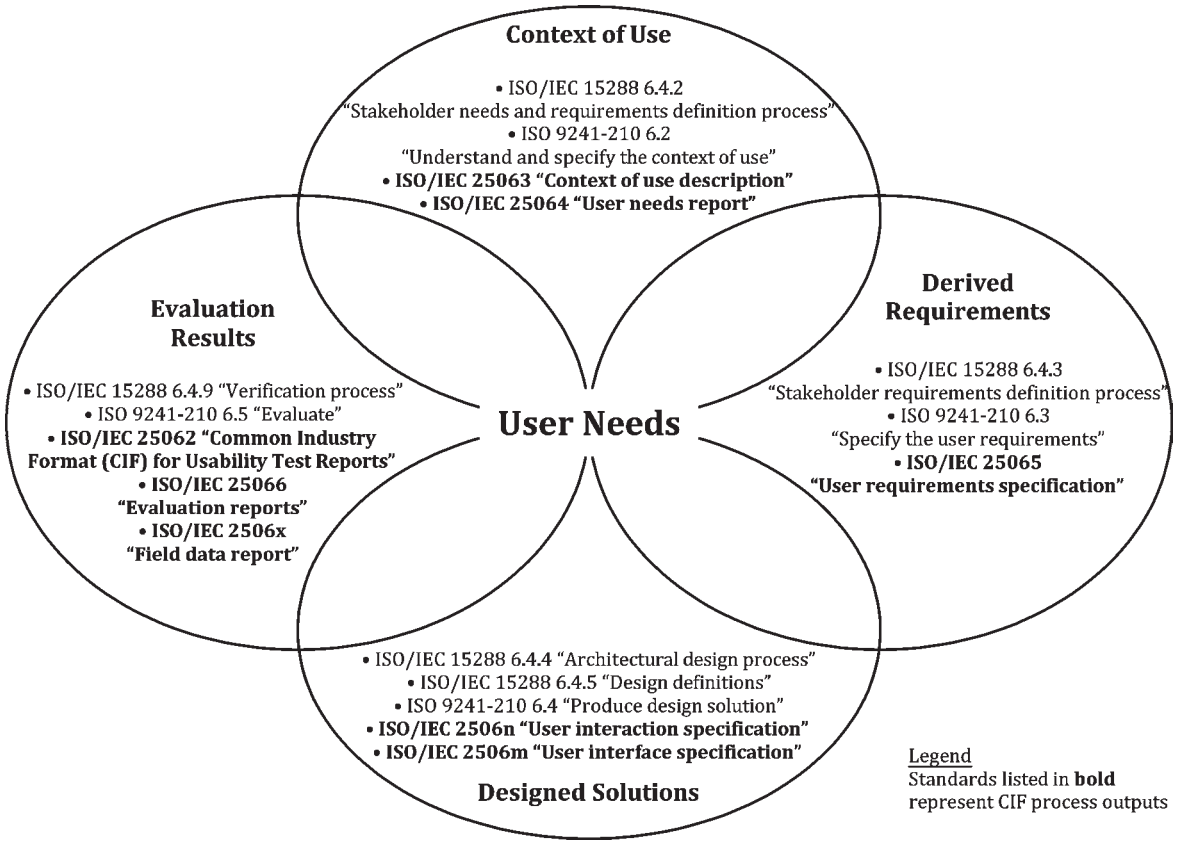

Figure 1 — Relationship of CIF documents to human-centred design in ISO 9241-210 and system lifecycle processes in ISO/IEC 15288

Figure 1 illustrates the interdependence of these information items with the human-centred design activities described in ISO 9241-210, as well as the corresponding System Life Cycle processes described in ISO/IEC 15288 [4].

The following discussion also serves as alternative text for the figure.

The figure depicts the activities as a set of intersecting circles. The circles overlap to represent that the activities are not separate, but rather overlapping in time and scope, and the outcome of each activity provides the input to one or more other activities. As each human-centred design activity can provide input to any other, no starting point, end point, or linear process is intended.

4) Withdrawn. Replaced with ISO/IEC/IEEE 15288:2015.

The human-centred design is composed of four interacting activities represented as overlapping circles in the diagram where User Needs are at the centre.

The first activity involves Context of Use. Human-centred design relies on user needs that are first identified during of the Context of Use analysis. User needs are documented in the User needs report (ISO/IEC 25064), which is an intermediate deliverable that links the Context of Use Description (ISO/IEC 25063) that contains information about the users, their tasks and the organizational and physical environment, to the user requirements. These items are developed during the Stakeholders requirements definition process described in ISO/IEC 15288.

The second activity involves Derived Requirements. The User requirements specification (ISO/IEC 25065) provides the basis for design and evaluation of interactive systems to meet the user needs. User requirements are developed in conjunction with and from part of the overall requirements specification of an interactive system.

The third activity involves Designed Solutions. The “Produce design solutions” activity focuses on designing user interaction that meets user requirements. This activity takes place during the Architectural Design, Implementation, and Integration processes described in ISO/IEC 15288 and produces the information items “User interaction specification” and the “User interface specification”.

The fourth activity involves Evaluation Results. The “Evaluate” activity starts at the earliest stages in the project, evaluating design concepts to obtain a better understanding of the user needs. Design solutions can be evaluated multiple times as the interactive system is being developed and can produce various types of evaluation reports and usability data such as that described in ISO/IEC 25062. These evaluations can support the ISO/IEC 15288 Validation Process that confirms that the system complies with the stakeholders’ requirements.

Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for Usability — Evaluation Report

1 Scope

This International Standard describes the Common Industry Format (CIF) for reporting usability evaluations. It provides a classification of evaluation approaches and the specifications for the content items (content elements) to be included in an evaluation report based on the selected evaluation approach(es). The intended users of the usability evaluation reports are identified, as well as the situations in which the usability evaluation report can be applied.

The usability evaluation reports in this International Standard are applicable to software and hardware systems, products or services used for predefined tasks (excluding generic products, such as a display screen or a keyboard). The content elements are intended to be used as part of system-level documentation resulting from development processes such as those in ISO 9241-210 and ISO/IEC JTC 1/SC 7 process standards.

The content elements for documenting evaluations can be integrated in any type of process model.

NOTE For the purpose of establishing process models, ISO/IEC TR 24774 and ISO/IEC 33020 specify the format and conformance requirements for process models, respectively. In addition, ISO/IEC 15289 defines the types and content of information items developed and used in process models for system and software lifecycle management. ISO/IEC 15504-5 and ISO/IEC 15504-6 (to be replaced by ISO/IEC 33060) define work products, including information items, for the purpose of process capability assessment. Process models and associated information items for human-centred design of interactive systems are contained in ISO/TR 18529 and ISO/TS 18152.

2 Conformance

An evaluation report conforms to this International Standard if it contains all the required content elements in Clause 5 that are applicable to the type(s) of evaluation, including:

— additional optional content elements that were selected to be part of the evaluation;

— the content elements for the conformity assessment (if used).

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

NOTE The CIF family of standards uses definitions that are consistent with the ISO 9241 series of standards, as this is the terminology that is normally used for this subject matter. In some cases, these definitions differ from those in ISO/IEC 25000.

3.1
accessibility
extent to which products, systems, services, environments and facilities can be used by people from a population with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use

Note 1 to entry: Context of use includes direct use or use supported by assistive technologies.

[SOURCE: ISO 26800:2011, 2.1; modified, Note 2 to entry deleted]

3.2
action
user behaviour that a system accepts as a request for a particular operation

[SOURCE: ISO/IEC TR 11580:2007, 2.3; modified, Example deleted]

3.3
conformity assessment
demonstration that specified requirements relating to a product, process, system, person or body are fulfilled

[SOURCE: ISO/IEC 17000:2004, 2.1; modified, Notes deleted]

3.4
context of use
users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used

[SOURCE: ISO 9241-11:1998, 3.5]

3.5
dialogue
interaction between a user and an interactive system as a sequence of user actions (inputs) and system responses (outputs) in order to achieve a goal

Note 1 to entry: User actions include not only entry of data but also navigational actions of the user.

Note 2 to entry: Dialogue refers to both the form (syntax) and the meaning (semantics) of interaction.

[SOURCE: ISO 9241-110:2006, 3.2]

3.6
effectiveness
accuracy and completeness with which users achieve specified goals

[SOURCE: ISO 9241-11:1998, 3.2]

3.7
efficiency
resources expended in relation to the accuracy and completeness with which users achieve goals

[SOURCE: ISO 9241-11:1998, 3.3]

3.8
goal
intended outcome

[SOURCE: ISO 9241-11:1998, 3.8]

3.9
information item
separately identifiable body of information that is produced and stored for human use during a system or software life cycle

[SOURCE: ISO/IEC/IEEE 15289:2011, 5.7]

3.10
inspection-based evaluation
evaluation based on the judgment of one or more evaluator(s) who examine or use a system to identify potential usability problems (including deviations from established criteria)

Note 1 to entry: The evaluators making the inspections typically are usability specialists but can also include end users and members of the design team.

Note 2 to entry: Established criteria typically include user requirements, usability guidelines in standards, design conventions contained in manufacturer guidelines and style guides, task models to be supported, as well as standardized principles.

Note 3 to entry: The evaluation can be conducted with or without the help of reference documents.

Note 4 to entry: Inspection-based evaluation is a generic term for methods that include but are not limited to heuristic evaluation, cognitive walkthroughs, standards inspection, pluralistic walkthroughs, and consistency inspections.

Note 5 to entry: Inspection-based evaluation can be conducted by machines in some cases, e.g. when consistency with required terminology is being evaluated. In this case, the machine represents the evaluator.

3.11
requirement
condition or capability that must be met or possessed by a system, system component, product, or service to satisfy an agreement, standard, specification, or other formally imposed documents

[SOURCE: ISO/IEC/IEEE 24765:2010, 3.2506, Clause 4.]

3.12
satisfaction
freedom from discomfort, and positive attitudes towards the use of the product

[SOURCE: ISO 9241-11:1998, 3.4]

3.13
stakeholder
individual or organization having a right, share, claim, or interest in a system or in its possession of characteristics that meet their needs and expectations

[SOURCE: ISO/IEC/IEEE 15288:2015, 4.1.44]

3.14
system
combination of interacting elements organized to achieve one or more stated purposes

Note 1 to entry: A system may be considered as a product or as the services it provides.

Note 2 to entry: In practice, the interpretation of its meaning is frequently clarified by the use of an associative noun, e.g. aircraft system. Alternatively, the word system may be substituted simply by a context dependent synonym, e.g. aircraft, though this may then obscure a system principles perspective.

[SOURCE: ISO/IEC/IEEE 15288:2015, 4.1.46; modified, Note 3 to entry deleted]

3.15
task
activities required to achieve a goal

Note 1 to entry: The term "task" is used here, as in ISO 9241-11:—[5)], in its widest sense, rather than in reference to the specifics of use of the dialogue system.

[SOURCE: ISO 9241-11:1998, 3.9; modified, Notes changed]

3.16
usability
extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use

Note 1 to entry: According to ISO/IEC 25010, "Usability can either be specified or measured as a product quality characteristic in terms of its sub-characteristics, or specified or measured directly by measures that are a subset of quality in use." The definition of usability in this International Standard is consistent with the second approach.

5) Under preparation.

[SOURCE: ISO 9241-210:2010, 2.13; modified, Notes changed]

3.17
usability defect
product attribute(s) that lead(s) to a mismatch between user intentions and/or user actions and the system attributes and behaviour

Note 1 to entry: Typical usability defects include the following:

— additional unnecessary steps not required as part of completing a task;

— misleading information;

— insufficient and/or poor information on the user interface;

— unexpected system responses;

— limitations in navigation;

— inefficient use error recovery mechanisms;

— physical characteristics of the user interface that are not suitable for the physical characteristics of the user.

Note 2 to entry: Deviations of product attributes of the object of evaluation from established criteria are also usability defects.

3.18
usability finding
identified usability defect and/or usability problem or positive usability-related attribute

3.19
usability problem
situation during use resulting in poor effectiveness, efficiency or satisfaction

3.20
use error
user action or lack of user action while using the interactive system that leads to a different result than that intended by the manufacturer or expected by the user

Note 1 to entry: Use error includes the inability of the user to complete a task.

Note 2 to entry: Use errors can result from a mismatch between the characteristics of the user, user interface, task, or use environment.

Note 3 to entry: Users might be aware or unaware that a use error has occurred.

Note 4 to entry: An unexpected physiological response of the patient is not by itself considered a use error.

Note 5 to entry: A malfunction of an interactive system that causes an unexpected result is not considered a use error.

[SOURCE: IEC 62366-1:2015, 3.21; modified, Medical device replaced by interactive system, Notes changed]

3.21
user
person who interacts with a system, product or service

Note 1 to entry: Users include people who operate a system, people who use the output provided by a system and people who conduct support tasks using the system (including maintenance and training).

Note 2 to entry: According to ISO/IEC 25010, User is defined as “individual or group that interacts with a system or benefits from a system during its utilization”.

Note 3 to entry: Primary and secondary users interact with a system, and primary and indirect users can benefit from a system. This definition includes a broader understanding of individuals and organisations that act as users.

[SOURCE: ISO 26800:2011, 2.10; modified, Notes changed]

3.22
user-based evaluation
evaluation that involves representative users performing tasks with the system to enable identification of usability problems and/or measurements of efficiency, effectiveness, user satisfaction or other user experiences

3.23
user experience
a person's perceptions and responses that result from the use and/or anticipated use of a product, system or service

Note 1 to entry: User experience includes all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use.

Note 2 to entry: User experience is a consequence of: brand image, presentation, functionality, system performance, interactive behaviour, and assistive capabilities of the interactive system, the user's internal and physical state resulting from prior experiences, attitudes, skills and personality, and the context of use.

Note 3 to entry: Usability, when interpreted from the perspective of the users' personal goals, can include the kind of perceptual and emotional aspects typically associated with user experience. Usability criteria can be used to assess aspects of user experience.

[SOURCE: ISO 9241-210:2010, 2.15]

3.24
user need
prerequisite identified as necessary for an user, or a set of users, to achieve an intended outcome, implied or stated within a specific context of use

EXAMPLE 1 A presenter (user) needs to know how much time is left (prerequisite) in order to complete the presentation in time (intended outcome) during a presentation with a fixed time limit (context of use).

EXAMPLE 2 An account manager (user) needs to know the number of invoices received and their amounts (prerequisite), in order to complete the daily accounting log (intended outcome) as part of monitoring the cash flow (context of use).

Note 1 to entry: A user need is independent of any proposed solution for that need.

Note 2 to entry: User needs are identified based on various approaches including interviews with users, observations, surveys, evaluations, expert analysis, etc.

Note 3 to entry: User needs often represent gaps (or discrepancies) between what should be and what is.

Note 4 to entry: User needs are transformed into user requirements considering the context of use, user priorities, trade-offs with other system requirements and constraints.

[SOURCE: ISO/IEC 25064:2013, 4.19]

3.25
user requirements
usage requirements
requirements for use that provide the basis for design and evaluation of interactive systems to meet identified user needs

Note 1 to entry: User requirements are derived from user needs, characteristics and capabilities in order to make use of the system in an effective, efficient, safe and satisfying manner.

Note 2 to entry: User requirements specify the extent to which user needs, characteristics and capabilities are to be met when using the system. They are not requirements on the users.

Note 3 to entry: In software-engineering terms, user requirements comprise both “functional” and “non-functional” requirements based on user needs and capabilities.

[SOURCE: ISO/IEC TR 25060:2010, 2.21]

4 Purpose and types of usability evaluations

4.1 Purpose of an evaluation

The content of a usability evaluation report varies based on the purpose of the evaluation. An evaluation could be performed to test whether specified user requirements have been implemented or to test whether specified accessibility recommendations have been implemented. Or an evaluation could be performed as the basis for a procurement decision. This International Standard describes the contents of usability evaluation reports produced for a broad range of usability evaluation objectives.

The purpose of ISO/IEC 25062 is to facilitate incorporation of usability as part of the procurement decision-making process for software to assist in judging if a product meets usability goals. Examples of decisions include purchasing, upgrading and automating. ISO/IEC 25062 is an example of a user observation report for performance measurement in accordance with Annex A. ISO/IEC 25062 provides a common format for human factors engineers and usability professionals in supplier companies to report the methods and results of usability tests to customer organizations. Since the procurement environment is the intended audience, ISO/IEC 25062 is more prescriptive in the format and the required elements.

4.2 Types of usability evaluations

Usability evaluation is a systematic process using one of the following types of evaluation approaches. The content of an evaluation report depends on the type of evaluation approach used.

The classification of evaluation approaches described below is used in Clause 2.

a) Inspection to identify usability defects and potential usability problems including:

— deviations of the object of evaluation from specified criteria such as user requirements, principles, design guidelines or established conventions;

— potential usability problems when attempting to complete one or more tasks with the object of evaluation.

b) Observation of users including:

— observing user behaviour to identify actual usability findings;

— measuring user performance and response (e.g. time taken to perform a task, number of use errors, skin conductance or eye pupil dilation).

NOTE 1 The observation of users can be carried out as an explicit usability test and/or conducted in a “real life” setting.

NOTE 2 The usability problems are either identified during the observation or are identified from subsequent analysis.

c) User surveys including:

— eliciting problems, opinions and impressions from users (qualitative user surveys);

— measuring level of user satisfaction or perception, e.g. rating scale values for satisfaction or for subjectively perceived effectiveness or efficiency (quantitative user surveys);

— other user reported data (e.g. data collected from an individual in conjunction with observation data).

NOTE 3 Collection of information about participants such as demographic data does not constitute a user survey.

A usability evaluation report contains information about one or more types of the evaluations listed above.

EXAMPLE 1 A usability test report describes problems encountered by users when carrying out tasks (type of information is "Observing user behaviour"). A quantitative usability test report based on ISO/IEC 25062 contains measures of effectiveness, efficiency and satisfaction (types of information are "Measuring user performance" and "User survey").

When reporting findings in usability evaluation reports, it is important to differentiate usability defects from their consequences. While usability defects are typically inappropriate attributes of the interactive system, their consequences describe the negative effect on the user that is either likely to occur or has been observed or reported.

EXAMPLE 2 A usability defect could be the fact, that within a web form, required entry fields are not marked as such. The consequences could be that users fail to fill in required entry fields and therefore make use errors repeatedly.

Usability evaluation content can be further categorized by the types of evaluation involved. Usability evaluations can be differentiated in terms of "inspection-based" versus "user-based". The following clauses introduce the general types of usability evaluation reports.

4.3 Assessing conformity of the object of evaluation against specified criteria

Evaluation report data can be used for different purposes. One purpose is, to show that the object of evaluation meets specified requirements, also referred to as conformance criteria. A conformity assessment of the object of evaluation against specified criteria is defined in ISO/IEC 17000 as a "demonstration that specified requirements relating to a product, process, system, person or body are fulfilled". Assessment of conformity consists of comparing the evaluation results with pre-defined conformance criteria. The conformance criteria can be defined within a project or by a third party (e.g. a regulatory body). A rigorous evaluation is required to produce data that can be used for a conformity assessment. When a conformity assessment is used, it shall be documented in conformance with the requirements of this International Standard.

NOTE A formal conformity assessment requires a defined "conformity assessment scheme". The formal scheme provides a) legal defensibility, b) evidence of contractual compliance, c) consistency of application and comparability of results across assessors and organizations. Conformity assessment schemes are implemented at an international, regional, national and sub-national level.

The conformity assessment can be included in a usability evaluation report or can be issued as a separate "conformity assessment report". Table 2 shows the different types of conformance criteria that can be specified as the basis for a conformity assessment. There can be various sets of specified conformance criteria for one conformity assessment, if the underlying evaluation consisted of more than one type of evaluation (e.g. inspection plus user observation plus user survey).

Table 2 — Conformance criteria used for conformity assessment and corresponding types of usability evaluation reports

Conformance criteria	Type of usability evaluation report
— Specified user requirements (e.g. "The user shall be able to sort flights by duration." or "The user shall be able to select alternative modes of input or output to carry out a task.") — Specified principles (e.g. "error tolerance") and guidelines (e.g. "Required entry fields shall be visually distinct from optional entry fields.") — Specified design conventions (e.g. "The edit-button is always at the top-right corner of the form.")	Inspection-based evaluation report
— Specified user requirements (e.g. "The user shall be able to detect that one or more patients need immediate attention.") — Specified user requirements for performance (e.g. "The user shall be able to complete the sales order within 60 seconds")	User observation report
— Specified scores for subjectively perceived effectiveness, efficiency, satisfaction and other measures perceived by users (e.g. 3,5 on a scale ranging from 1 (min) to 5 (max)) — Specified attributes for reported experiences (e.g. "If any of the reported usability problems is judged as unacceptable then the object of evaluation fails the conformity assessment.")	User survey report

Principles and guidelines that can be used as conformance criteria are published in various sources including the ISO 9241 series. These principles and guidelines often apply across operating systems and development environments, e.g. "Colour should not be used as the only means to code information." or "Required entry fields should be visually distinct from optional entry fields." User-interface related recommendations can be found in the ISO 9241 series of standards:

— ISO 9241-12 — Presentation of information;

— ISO 9241-13 — User guidance;

— ISO 9241-14 — Menu dialogues;

— ISO 9241-15 — Command dialogues;

— ISO 9241-16 — Direct manipulation dialogues;

— ISO 9241-20 — Accessibility guidelines for information/communication technology (ICT) equipment and services;

— ISO 9241-110 — Dialogue principles;

— ISO 9241-129 — Guidance on software individualization;

— ISO 9241-143 — Forms;

— ISO 9241-151 — Guidance on World Wide Web user interfaces;

— ISO 9241-171 — Guidance on software accessibility;

— ISO 9241-303 — Requirements for electronic visual displays;

— ISO 9241-400 — Principles and requirements for physical input devices;

— ISO 9241-410 — Design criteria for physical input devices;

— ISO 9241-920 — Guidance on tactile and haptic interactions.

Established conventions that can also be used as conformance criteria typically include rules published by suppliers of operating systems (e.g. “Windows”, “Mac OS”, iOS”, “Android”) and development environments (e.g. “.NET” or “Java”).

EXAMPLE An example for an established convention is “a dialog box always has an “OK” and “Cancel” button at the bottom right corner of the dialogue box”.

5 Content elements of usability evaluation reports

5.1 Selecting content elements

The following clauses describe the content elements that can be included in a usability evaluation report. The content elements are described in subclauses organized on the basis of sections that can be included in an evaluation report.

Depending on the purpose of the evaluation, a usability evaluation report can include the following sections:

— Executive Summary;

— Description of the object of evaluation;

— Purpose of evaluation;

— Method;

— Procedure;

— Results;

— Interpretation of results and recommendations (optional).

Within each subclause, the required, recommended and permitted content elements for each type of evaluation are indicated within a table at the end of each subclause. Each content element is specified as: mandatory, i.e. required (“shall”), recommended (“should”) or permitted (“may”) for each type of evaluation (i.e. inspection, user observation and user survey). Requirements describe elements that are essential in all situations. Recommendations are also important but they might not apply in all situations.

The content elements for each section of an evaluation report are determined by the type(s) of evaluation to be conducted. Furthermore, there are elements, that are always required and conditional elements, that can be selected for the evaluation, if used (e.g. statistical analysis or provided recommendations) and/or applicable (e.g. parts of the object that were evaluated or measures used in evaluation).

Evaluations often contain more than one type of evaluation (e.g. user observation and subsequent user survey). As a result, the evaluation report would include the content elements for both types of evaluations.

The order in which the sections and the elements within it are introduced does not prescribe a required order for a usability evaluation report. Furthermore, the grouping of the content elements themselves can be defined by the author of the report (e.g. combining information such as methods and procedures into one section of the evaluation report).

The evaluation report should provide sufficient information to determine the appropriateness of the evaluation and to assess the validity of the results.

NOTE For user observation, the context of use for evaluation needs to reproduce the key aspects of a subset of the context of use in order for evaluation results to be valid.

Annex A contains a table that gives an overview of all required and recommended content elements for each type of evaluation.

The following subclauses enumerate the content elements for a usability evaluation report. The report sections described in the following clauses refer to all three types of evaluation (inspection-based evaluation, user observation and user survey).

5.2 Description of the content elements for each type of evaluation

5.2.1 Executive summary (if used)

This section of the usability evaluation report provides a concise overview of the evaluation. The intent of this section is to provide information for those who might not read the technical body of the report.

An executive summary can include:

a) Name and description of the object of evaluation;

b) Summary of method(s) and the procedure;

c) Summary of results including key findings, related conclusions and recommendations (if applicable).

Table 3 specifies the required and recommended items for each type of evaluation.

Table 3 — Executive Summary

Type of evaluation: **Content elements to be included in report:**	**Inspection**	**User observation**		**User survey**
		Observing user behaviour	**Measuring user performance and response**	
a) Name and description of the object of evaluation	shall	shall	shall	shall
b) Summary of method(s) and the procedure	shall	shall	shall	shall
c) Summary of results including key findings, related conclusions and recommendations (if applicable)	shall	shall	shall	shall

5.2.2 Description of the object of evaluation

This section of the usability evaluation report identifies the entity, which was actually evaluated.

NOTE Examples of objects of evaluation include concepts, user interface prototypes, functioning software systems, hardware products, or components of a product or a service.

Information about the object of evaluation can include:

a) Formal name and release or version;

b) Parts of the object that were evaluated (if applicable);

c) User groups for which the object is intended;

d) Brief description of the object and its purpose;

e) Intended context of use;

f) Prior usability evaluation report summaries (if applicable);

g) Expected impact (e.g. on performance, safety, finances) of the object;

h) Citations to market research for the object.

The context of use for the object of evaluation needs to be described. Further guidance on the description of the context of use is given in ISO/IEC 25063. Each of the four components of the context

of use (users, tasks, equipment, environment) is not always applicable for every type of evaluation (for example, tasks are not always used for inspection-based evaluations).

Table 4 specifies the required and recommended and permitted items for each type of evaluation.

Table 4 — Description of the object of evaluation

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Formal name and release or version	shall	shall	shall	shall
b) Parts of the object that were evaluated (if applicable)	shall	shall	shall	shall
c) User groups for which the object is intended	should	should	should	should
d) Brief description of the object and its purpose	should	should	should	should
e) Intended context of use	should	should	should	should
f) Prior usability evaluation report summaries (if applicable)	should	should	should	should
g) Expected impact of the object	may	may	may	may
h) Citations to market research for the object	may	may	may	may

5.2.3 Purpose of the evaluation

This section of the usability evaluation report identifies the reasons for which the evaluation was conducted and which parts of the object were evaluated and why.

a) Description of the purpose

NOTE 1 Purposes for an evaluation can include:

— improving design by providing feedback into the design process;

— identifying usability defects and usability problems;

— confirming/eliciting user requirements;

— confirming assumptions;

— testing concepts;

— measuring the level of usability (i.e. effectiveness and/or efficiency and/or user satisfaction);

— establishing benchmarks;

— assessing whether a product, system or service meets specific conformance criteria/ acceptance criteria;

— identifying strengths and weaknesses of a product, system or service;

— identifying the consequences that could arise from poor usability;

— resolving disputes between users and/ or stakeholders;

— identifying whether a product, system or service is accessible;

— acquiring a certification, e.g.

  — to pass an internal quality gate;

  — to pass a certification of a certification body.

b) Functions and components evaluated (if applicable)

NOTE 2 It is not necessary to describe the functions and components, if all the functions and components were evaluated.

c) Reasons for only evaluating a subset of the object (if applicable)

NOTE 3 It is not necessary to describe the reasons why only a part of the object was evaluated, if all the functions and components were evaluated.

Table 5 specifies the required and recommended items for each type of evaluation.

Table 5 — Purpose of the evaluation

Type of evaluation: **Content elements to be included in report:**	**Inspection**	**User observation**		**User survey**
		Observing user behaviour	**Measuring user performance and response**	
a) Description of the purpose	shall	shall	shall	shall
b) Functions and components evaluated (if applicable)	shall	shall	shall	shall
c) Reasons for only evaluating a subset of the object (if applicable)	shall	shall	shall	shall

5.2.4 Method

5.2.4.1 General

This section of the evaluation report describes how the evaluation was conducted. The goal of the description is to provide enough information to determine the appropriateness of the method and to assess the validity of the results, as well as enabling replication.

a) Type(s) of evaluation used

Usability evaluation reports can include data based on one or more than one type of evaluation (see 4.1). The usability evaluation report states which type(s) of evaluation have been used.

NOTE Types of evaluations are inspection-based evaluation, observing user behaviour, measuring user performance and user survey.

b) Sufficient information to replicate the procedure used during the evaluation

Table 6 specifies the required items for each type of evaluation.

Table 6 — General

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Type(s) of evaluation used	shall	shall	shall	shall
b) Sufficient information to replicate the evaluation procedure used during the evaluation	shall	shall	shall	shall

5.2.4.2 Evaluators/participants

This section of the evaluation report provides information about the people taking part in the usability evaluation. Evaluators are the people who run the evaluation, and include people who carry out inspections. Participants are people who are actual or potential users of the object of evaluation, who take part in observational studies where their behaviour and/or task performance is monitored. The people who provide survey data are also participants.

Information about the evaluators and participants enables readers of a report to judge whether the information presented is applicable to their own circumstances.

a) Total number of evaluators/participants

This element reports the total number of evaluators or participants.

b) Segmentation of test participants or evaluators/inspectors (if more than one segment)

When observing user behaviour, segmentation of test participants into groups based on their characteristics can be used as an experimental variable (see 5.2.5.1). Segmentation of evaluators enables the results produced by different categories of evaluator to be compared.

EXAMPLE 1

— evaluator/inspector with domain expertise;

— evaluator/inspector with usability expertise;

— evaluator/inspector representing the users;

— segmentation of test participants (if more than one).

EXAMPLE 2

— infrequent users versus habitual users.

c) Key Characteristics of test participants or users considered for inspection

Key characteristics of test participants characterize attributes of the intended user population that are relevant to the validity of the evaluation.

NOTE 1 Key characteristics can include:

— demographics that are used to identify intended user groups of specific interest, for example, age;

— task-related characteristics, for example, training, skill level and established behaviours;

— physical and sensory characteristics, for example body dimensions, strength, vision and hearing;

— psychological and social characteristics, for example reading age, habits, language and culture;

— social and organizational characteristics, for example profession or job title, resistance to change and a risk taking culture;

— user group membership (i.e. the groups that the test participant represents for this evaluation, e.g. smart-phone users, land-line phone users).

d) Differences between sample and the user population (if applicable)

This element describes any differences between the participant sample and the actual user population. In particular differences in key characteristics are described.

EXAMPLE 3 Actual users might attend a training course whereas test subjects were untrained.

e) Table of participants by characteristics

Tables help summarize and improve the readability of the key characteristics of participants for the reader.

NOTE 2 A table can include:

— participants (row) by characteristics (columns), of key characteristics such as computing experience, age, gender, abilities.

Table 7 specifies the required and recommended items for each type of evaluation.

Table 7 — Evaluators/participants

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Total number of evaluators/ participants	shall	shall	shall	shall
b) Segmentation of test participants or evaluators/ inspectors (if more than one segment)	shall	shall	shall	shall
c) Key characteristics of test participants or users considered for inspection	shall	shall	shall	shall
d) Differences between sample and the user population (if applicable)	N/A	shall	shall	shall
e) Table of participants by characteristics	N/A	shall	shall	shall

5.2.4.3 Tasks (if used in the evaluation)

This section of the evaluation report describes the tasks used for evaluation. When observing user behaviour, measuring user performance or gathering survey data, typically tasks are specified. Inspections can also be tasked-based, but not all inspections are task-based. For inspections, the usability evaluation report shall explicitly state whether or not tasks are specified.

The information if tasks are used can include:

a) Tasks used for evaluation

The tasks used for evaluation are expressed in terms of the title and intended outcomes that people are expected to achieve without referencing any specific means of achieving them.

b) Task scenarios for each task

A task scenario is the information provided to the participants including any materials handed out.

c) Selection criteria for the tasks

The selection criteria for the tasks explain why the selected tasks were deemed to be important for the evaluation.

EXAMPLE 1 The most frequent tasks for each selected user group.

EXAMPLE 2 Tasks that give rise to the greatest potential risk.

d) Source of selected tasks

The source of the selected tasks explains what the tasks are based on.

EXAMPLE 3 Observation of customers using similar products, product marketing specifications, discussion with users or design team.

e) Task data given to participants and/or inspectors (if applicable).

EXAMPLE 4 Data to be processed.

EXAMPLE 5 Recorded simulation of the customer request.

f) Criteria for task completion or task abandonment for each task.

These are the criteria for terminating the task, either finishing the task or quitting.

EXAMPLE 6 After more than 30 min, the task is terminated.

EXAMPLE 7 After three unsuccessful attempts, the task is terminated.

EXAMPLE 8 As soon as the user believes the task has been completed.

Table 8 specifies the required and recommended items for each type of evaluation.

Table 8 — Tasks (if used in the evaluation)

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Tasks used for evaluation	shall	shall	shall	shall
b) Task scenarios for each task	shall	shall	shall	shall
c) Selection criteria for the tasks	shall	shall	shall	shall
d) Source of selected tasks	shall	shall	shall	shall
e) Task data given to participants and/ or inspectors (if applicable)	shall	shall	shall	shall
f) Criteria for task completion and task abandonment for each task	N/A	may	shall	N/A

5.2.4.4 Evaluation environment

a) Physical environment and facilities

This section of the evaluation report describes the information related to the physical environment and facilities.

NOTE 1 This can include a description of the setting and type of space in which the evaluation was conducted.

NOTE 2 This is especially important in situations in which the context of evaluation is different to the intended context of use:

EXAMPLE 1 Usability lab, cubicle office, meeting room, home office, home family room, manufacturing floor, remote usability testing using video and audio conferencing and desktop sharing, etc.

— any relevant features of the setting or circumstances that could affect the results.

EXAMPLE 2 Video and audio recording equipment, one-way mirrors, or automatic data collection equipment

— a description of the physical environment.

EXAMPLE 3 Dark environment (lighting) in a radiological screening room

EXAMPLE 4 The user's choice of location or environment (e.g. home, car, office, etc.) is reported for remote testing

b) Technical environment (if applicable)

This element describes both the software and computing environment.

NOTE 3 This can include:

— computer configuration, including model, operating system version, required libraries or settings, if browser-based, browser name and version; relevant plug-in names and versions;

— any other display devices, audio devices, and input devices.

c) Evaluation administration tools (if used)

This element describes any hardware or software used to control the evaluation or to record data.

NOTE 4 This can include:

— a description or specification of a standard questionnaire;

— any hardware or software used to control the test or to record data.

d) Evaluation administrators (if applicable)

This element identifies the number of evaluation facilitators/administrators and their roles and responsibilities.

Table 9 specifies the required and recommended items for each type of evaluation.

Table 9 — Evaluation environment

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Physical environment and facilities	N/A	shall	shall	may
b) Technical environment (if applicable)	shall	shall	shall	should
c) Evaluation administration tools (if used)	should	should	should	should
d) Evaluation administrators (if applicable)	may	should	should	should

5.2.5 Procedure

5.2.5.1 Design of the evaluation

This section of the evaluation report details each step in the execution of the usability evaluation. This section summarizes what one did and how one did it. It identifies the type of evaluation, the specific experimental manipulations, as well as instructions to the participants.

a) Description of the evaluation design

This element describes the type of evaluation performed and the experimental design of the evaluation, the plan for assigning experimental conditions to participants including the specific experimental manipulations if applicable.

NOTE 1 Experimental manipulations can include:

— randomization;

— counterbalancing;

— other control features.

b) Independent variables (if applicable)

This element describes those variables that are manipulated by the evaluator (independent variables).

EXAMPLE 1 Different levels of training, age of participants, noise levels, lighting, which prototype participants experience.

c) Predefined criteria (for inspection or observation) (if used)

This element describes the criteria used for inspection or observation.

NOTE 2 This can include:

— for inspection:
  — principles;
  — guidelines;
  — established conventions;
— for inspection or user observation:
  — user requirements.

EXAMPLE 2 A specified principle to be inspected against is “error tolerance”.

EXAMPLE 3 A specified guideline to be inspected against is “required entry fields shall be visually distinct from optional entry fields.”

EXAMPLE 4 A specified accessibility guideline to be inspected against is “Whenever moving, blinking, scrolling, or auto-updating information is presented, software shall enable the user to pause or stop the presentation, except for simple progress indicators.”

EXAMPLE 5 A specified user requirement to be inspected against is “The user shall be able to sort flights by duration”.

EXAMPLE 6 A specified design convention to be inspected against is “The edit-button is always at the top-right corner of the form”.

d) Measures used in evaluation (if applicable)

This element describes the measures for which data were recorded for each set of conditions.

EXAMPLE 7 Number of use errors participants make during a task.

e) Operational definitions of criteria or measures (if applicable)

This element describes what constitutes a criteria or measure.

EXAMPLE 8 What constitutes a participant use error: “An incorrect navigational choice”.

f) Interaction between individuals taking part in each evaluation session (if applicable)

This element describes the allowed interactions between individuals taking part in the evaluation.

NOTE 3 This can include:

— number and roles of testing staff and participants who will interact during the evaluation session;
— number and roles of participants and if and how they will interact with each other during the evaluation session.

g) Other individuals present in evaluations (if applicable)

This element identifies any other individuals expected to be present during the evaluation, if applicable.

h) General instructions given to the participants

This element describes the general instructions given to the participants of the evaluation.

NOTE 4 This can include:

— the actual instructions given to the participants (here or in an Appendix);

— instructions on how participants were to interact with any other persons present, including how they were to ask for assistance and interact with other participants, if applicable.

i) Specific instructions on tasks (if applicable)

This element describes the specific task instructions including time limits for completing the tasks given to participants or evaluator(s).

NOTE 5 This can include:

— task instruction summary;

— time limits per task.

j) Sequence of activities for conducting the evaluation

This element describes the sequence of organizational activities for conducting each session within the evaluation from greeting the participants to dismissing them.

NOTE 6 This can include:

— steps followed to execute the test sessions and record data;

— details of nondisclosure agreements, form completion, warm-ups, pre-task training, and debriefing;

— whether participants were paid or otherwise compensated;

— verification that the participants knew and understood their rights as human subjects.

Table 10 specifies the required and recommended items for each type of evaluation.

Table 10 — Design of the evaluation

Type of evaluation: **Content elements to be included in report:**	**Inspection**	**User observation**		**User survey**
		Observing user behaviour	**Measuring user performance and response**	
a) Description of the evaluation design	shall	shall	shall	shall
b) Independent variables (if applicable)	N/A	shall	shall	shall
c) Predefined criteria (for inspection or observation) (if used)	shall	shall	N/A	N/A
d) Measures used in evaluation (if applicable)	N/A	may	shall	shall
e) Operational definitions of criteria or measures (if applicable)	shall	shall	shall	shall
f) Interaction between individuals taking part in each evaluation session (if applicable)	N/A	shall	shall	should
g) Other individuals present in evaluations (if applicable)	N/A	should	should	should
h) General instructions given to the participants	N/A	shall	shall	shall
i) Specific instructions on tasks (if applicable)	should	shall	shall	N/A
j) Sequence of activities for conducting the evaluation	N/A	should	should	may

5.2.5.2 Data to be collected

This subclause describes the elements of the procedure that describe the data to be collected during the evaluation.

a) Usability defects in terms of deviations from predefined criteria (if criteria are used)

This element identifies that deviations from predefined criteria (see 5.2.5.1) will be collected during the evaluation.

NOTE 1 Deviations include all attributes of the interaction with the object of evaluation that deviate from criteria and are expected to cause usability problems.

EXAMPLE 1 A dialogue box that does not have a “Cancel” button.

EXAMPLE 2 In a context of use where there is a user requirement, that critical controls are readily identifiable (criterion), no user has found the emergency shutdown button (deviation).

b) Observed user behaviour

This element identifies all the types of observed user behaviour that are to be collected.

These data can be used in order to identify usability findings in the interaction with the object of evaluation.

NOTE 2 User observation data can include situations in which

— users don’t know how to proceed with the task,

— use errors occur,

— users communicate frustration,

— users engage in positive behaviours,

— users exhibit discomfort,

— evidence that one or more specific user requirements are met (or not met).

c) Observed performance data

This element identifies all the types of performance data relating to effectiveness and efficiency that are to be collected.

Performance data are a specific case of observation data where numerical values are obtained with a focus on measurement.

NOTE 3 Performance data measures can include:

— accuracy and completeness of task results (effectiveness) (if applicable);

— task completion;

— time taken on task;

— use errors and frequency of occurrence;

— number of mouse clicks, touch events or gestures;

— number of key strokes;

— distance moved on screen with pointing device (e.g. mouse);

— eye tracking paths;

— behavioural data (e.g. emotional, fidgeting, level of attention);

— physiological data (e.g. skin conductance, blood pressure).

d) User-reported qualitative data

This element identifies the type(s) of user-reported qualitative data to be collected and the data instrument that will be used.

User-reported qualitative data are statements made by users on their experience with the object of evaluation.

NOTE 4 A questionnaire is a typical instrument for collecting user-reported data. A questionnaire used to collect qualitative data uses open-ended questions.

NOTE 5 User-reported qualitative data can include:

— experienced problems;

— positive experiences;

— how the users use the object of evaluation;

— expectations;

— concerns;

— suggestions.

NOTE 6 User-reported problems can be accompanied by subjective severity ratings.

e) User-reported quantitative data

This element identifies the type(s) of user-reported quantitative data to be collected and the data instrument that will be used to collect.

User-reported quantitative data are experiences rated on a predefined scale.

NOTE 7 A questionnaire is a typical instrument for collecting user-reported data. A questionnaire used to collect quantitative data uses closed questions typically with an associated rating scale.

NOTE 8 User-reported quantitative data can include subjective ratings of the object of evaluation in terms of

— satisfaction,

— comfort,

— trustworthiness,

— attitude,

— appeal,

— effort,

— perceived effectiveness,

— perceived efficiency.

Table 11 specifies the required and recommended items for each type of evaluation.

Table 11 — Data to be collected

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Usability defects in terms of deviations from predefined criteria (if criteria are used)	shall	N/A	N/A	N/A
b) Observed user behaviour	N/A	shall	should	N/A
c) Observed performance data	N/A	may	shall	N/A
d) User-reported qualitative data	N/A	may	may	shall
e) User-reported quantitative data	N/A	N/A	N/A	shall

5.2.6 Results

5.2.6.1 Data analysis

This section of the usability evaluation report describes the data collected and the statistical or data analytic treatment used. Sufficient detail is reported to justify the conclusions. This section presents the results, a discussion of the results and the implication or inferences of the results.

a) Approach used for the analysis of observed, measured or collected data

This element describes in which way the observed, measured or collected data were analysed.

b) Differences in planned and collected data (if applicable)

This element describes the differences between the data that was planned to be collected and the data that was actually collected if applicable.

c) Portion of data used in the analysis (if applicable)

This element describes the portion of the gathered data that was actually used for the analysis.

EXAMPLE 1 How missing data was treated. How data was treated with respect to exclusion of outliers.

d) Data scoring (if used)

This element describes the mapping between the data values that were collected and the values used in the subsequent analysis.

EXAMPLE 2 How use errors were categorized. How actual ages map to age ranges. How assisted use errors are mapped to a set of values.

e) Data reduction

This element identifies the method used to generate summaries of the raw data.

EXAMPLE 3 Which measure of central tendency was used (e.g. mean or mode). How variation was measured (e.g. standard deviation or range).

EXAMPLE 4 Systematic characterization of open ended responses.

f) Statistical analyses (if used)

This element identifies and describes the statistical analyses used to analyse the data including the statistical procedure.

EXAMPLE 5 How groups were compared (e.g. t-test. F-test).

For data that are calculated as means calculate the standard deviation and the standard error of the mean.

Table 12 specifies the required and recommended items for each type of evaluation.

Table 12 — Data analysis

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Approach used for the analysis of observed, measured or collected data	shall	shall	shall	shall
b) Differences in planned and collected data (if applicable)	shall	shall	shall	shall
c) Portion of data used in the analysis (if applicable)	shall	shall	shall	shall
d) Data scoring (if used)	N/A	shall	shall	shall
e) Data reduction	N/A	shall	shall	shall
f) Statistical analyses (if used)	N/A	shall	shall	shall

5.2.6.2 Presentation of the results

This section of the evaluation report shows the presentation of the data collected based on the data analysis. Individual scores or raw data are not included. For data that are reported as means include the standard deviation and the standard error of the mean ("confidence interval") see 5.2.6.1. Tables and graphs can often be used to present complex results in a concise manner. Both tables and various graphical formats are effective in describing usability data at a glance. Bar graphs are useful for describing subjective data such as that gleaned from Likert scales. A variety of plots can be used effectively to show comparisons of expert benchmark times for the object of evaluation vs. the mean participant performance time. If necessary, the details of the interpretation are described in "findings and recommendations".

The results summarize:

a) Usability defects in terms of the deviations of attributes of the object of evaluation from established criteria

This element summarizes the usability defects in terms of the deviations of attributes of the object of evaluation from principles, guidelines and established conventions or specified user requirements.

NOTE 1 A mapping of deviations of attributes of the object of evaluation to specified criteria can be used to present the results.

b) Potential usability problems that are likely to arise from identified usability defects

This element summarizes the potential usability problems (and the associated rationales for the prediction) that result from the identified usability defects, i.e. deviations of attributes of the object of evaluation from principles, guidelines and established conventions or specified user requirements.

c) Usability findings identified during observations

This element summarizes the usability findings identified during observations.

d) Performance data based on measurements

This element summarizes the collected measures that characterize the performance results per task or task group.

NOTE 2 Performance data can be accumulated while observing user behaviour. These can include:

— accuracy and completeness of task results (effectiveness);

— task completion rate;

— time on task;

— efficiency;

— use errors and frequency of occurrence;

— number of assists;

— number of mouse clicks, touch events and gestures;

— number of key strokes;

— distance moved on screen with pointing device (e.g. mouse);

— eye tracking paths;

— psychological data (e.g. emotional, fidgeting, level of attention);

— physiological data (e.g. skin conductance, blood pressure).

e) Problems, opinions and impressions reported by users

This element summarizes the problems, opinions and impressions reported by users.

NOTE 3 Users can report problems, opinions and impressions spontaneously while being observed.

f) Measured level of user satisfaction or perception.

This element summarizes the measured level of user satisfaction or perception.

Table 13 specifies the required and recommended items for each type of evaluation.

Table 13 — Presentation of the results

Type of evaluation: Content elements to be included in report	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Usability defects in terms of the deviations of attributes of the object of evaluation from predefined criteria (if criteria are used)	shall	may	may	may
b) Potential usability problems that are likely to arise from identified usability defects	shall	N/A	N/A	N/A
c) Usability findings identified during observations	N/A	shall	should	N/A
d) Performance data based on measurements	N/A	may	shall	N/A
e) Problems, opinions and impressions reported by users	N/A	should	may	shall
f) Measured level of user satisfaction or perception	N/A	N/A	should	should

5.2.7 Interpretation of results and recommendations

This section of the usability evaluation report provides interpretation of results and recommendations, which help to identify the issues to be examined in detail.

NOTE In some cases, only the key interpretations and recommendations are included.

a) Interpretation of results

This element provides conclusions based on the interpretation of the results.

b) Recommendations

This element provides a set of recommendations for the improvement of the object of evaluation based on the evaluation results and their interpretation.

Table 14 specifies the required and recommended items for each type of evaluation.

Table 14 — Interpretation of results and recommendations

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Interpretation of results	should	should	should	should
b) Recommendations	should	should	should	should

5.2.8 Additional content for conformity assessment (as part of a usability evaluation report)

This section of the usability evaluation report provides conformity assessment. Usability evaluation reports for conformity assessment include additional content elements.

a) Conformity assessment scheme (if used)

This element describes the conformity assessment scheme (title, version).

NOTE Conformity assessment schemes can exist at international, regional, national or sub-national level.

b) Conformance criteria

This element describes the conformance criteria used.

c) Statement whether all conformance criteria have been met

This element provides a statement and describes whether all conformance criteria have been met.

d) Evidence showing why conformance criteria were not met (identified nonconformities)

This element provides the findings that show why specific conformance criteria have not been met (identified nonconformities).

Table 15 specifies the required and recommended items for each type of evaluation.

Table 15 — Additional content for conformity assessment (if used)

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
a) Conformity assessment scheme (if used)	shall	shall	shall	shall
b) Conformance criteria	shall	shall	shall	shall
c) Statement whether all conformance criteria have been met	shall	shall	shall	shall
d) Evidence showing why conformance criteria were not met (identified nonconformities)	shall	shall	shall	shall

Annex A
(normative)

Overview on required and recommended content elements for each type of evaluation

Table A.1 — Overview on required and recommended content elements for each type of evaluation

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
5.2.1 Executive summary				
a) Name and description of the object of evaluation	shall	shall	shall	shall
b) Summary of method(s) and the procedure	shall	shall	shall	shall
c) Summary of results including key findings, related conclusions and recommendations (if applicable)	shall	shall	shall	shall
5.2.2 Description of the object of evaluation				
a) Formal name and release or version	shall	shall	shall	shall
b) Parts of the object that were evaluated (if applicable)	shall	shall	shall	shall
c) User groups for which the object is intended	should	should	should	should
d) Brief description of the object and its purpose	should	should	should	should
e) Intended context of use	should	should	should	should
f) Prior usability evaluation report summaries (if applicable)	should	should	should	should
g) Expected impact of the object	may	may	may	may
h) Citations to market research for the object	may	may	may	may
5.2.3 Purpose of the evaluation				
a) Description of the purpose	shall	shall	shall	shall
b) Functions and components evaluated (if applicable)	shall	shall	shall	shall
c) Reasons for only evaluating a subset of the object (if applicable)	shall	shall	shall	shall
5.2.4 Method				
5.2.4.1 General				
a) Type(s) of evaluation used	shall	shall	shall	shall
b) Sufficient information to replicate the evaluation procedure used during the evaluation	shall	shall	shall	shall
5.2.4.2 Evaluators/ participants				

Table A.1 *(continued)*

Type of evaluation: **Content elements to be included in report:**	**Inspection**	**User observation**		**User survey**
		Observing user behaviour	**Measuring user performance and response**	
a) Total number of evaluators/ participants	shall	shall	shall	shall
b) Segmentation of test participants or evaluators/inspectors (if more than one segment)	shall	shall	shall	shall
c) Key characteristics of test participants or users considered for inspection	shall	shall	shall	shall
d) Differences between sample and the user population (if applicable)	N/A	shall	shall	shall
e) Table of participants by characteristics	N/A	shall	shall	shall
5.2.4.3 Tasks (if used in the evaluation)				
a) Tasks used for evaluation	shall	shall	shall	shall
b) Task scenarios for each task	shall	shall	shall	shall
c) Selection criteria for the tasks	shall	shall	shall	shall
d) Source of selected tasks	shall	shall	shall	shall
e) Task data given to participants and/ or inspectors (if applicable)	shall	shall	shall	shall
f) Criteria for task completion and task abandonment for each task	N/A	may	shall	N/A
5.2.4.4 Evaluation environment				
a) Physical environment and facilities	N/A	shall	shall	may
b) Technical environment (if applicable)	shall	shall	shall	should
c) Evaluation administration tools (if used)	should	should	should	should
d) Evaluation administrators (if applicable)	may	should	should	should
5.2.5 Procedure				
5.2.5.1 Design of the evaluation				
a) Description of the evaluation design	shall	shall	shall	shall
b) Independent variables (if applicable)	N/A	shall	shall	shall
c) Predefined criteria (for inspection or observation) (if used)	shall	shall	N/A	N/A
d) Measures used in evaluation (if applicable)	N/A	may	shall	shall
e) Operational definitions of criteria/ measures (if applicable)	shall	shall	shall	shall
f) Interaction between individuals taking part in each evaluation session (if applicable)	N/A	shall	shall	should
g) Other individuals present in evaluations (if applicable)	N/A	should	should	should
h) General instructions given to the participants	N/A	shall	shall	shall
i) Specific instructions on tasks (if applicable)	should	shall	shall	N/A

Table A.1 *(continued)*

Type of evaluation: Content elements to be included in report:	Inspection	User observation		User survey
		Observing user behaviour	Measuring user performance and response	
j) Sequence of activities for conducting the evaluation	N/A	should	should	may
5.2.5.2 Data to be collected				
a) Usability defects in terms of deviations from predefined criteria (if criteria are used)	shall	N/A	N/A	N/A
b) Observed user behaviour	N/A	shall	should	N/A
c) Observed performance data	N/A	may	shall	N/A
d) User-reported qualitative data	N/A	may	may	shall
e) User-reported quantitative data	N/A	N/A	N/A	shall
5.2.6 Results				
5.2.6.1 Data analysis				
a) Approach used for the analysis of observed, measured or collected data	shall	shall	shall	shall
b) Differences in planned and collected data (if applicable)	shall	shall	shall	shall
c) Portion of data used in the analysis (if applicable)	shall	shall	shall	shall
d) Data scoring (if used)	N/A	shall	shall	shall
e) Data reduction	N/A	shall	shall	shall
f) Statistical analyses (if used)	N/A	shall	shall	shall
5.2.6.2 Presentation of the results				
a) Usability defects in terms of deviations of attributes of the object of evaluation from predefined criteria (if criteria are used)	shall	may	may	may
b) Potential usability problems likely to arise from identified usability defects	shall	N/A	N/A	N/A
c) Usability findings identified during observations	N/A	shall	should	N/A
d) Performance data based on measurements	N/A	may	shall	N/A
e) Problems, opinions and impressions reported by users	N/A	should	may	shall
f) Measured level of user satisfaction or perception	N/A	N/A	should	should
5.2.7 Interpretation of results and recommendations				
a) Interpretation of results	should	should	should	should
b) Recommendations	should	should	should	should
5.2.8 Additional content for conformity assessment (if used)				
a) Conformity assessment scheme, if used	shall	shall	shall	shall
b) Conformance criteria	shall	shall	shall	shall
c) Statement, whether all conformance criteria have been met	shall	shall	shall	shall
d) Evidence showing why conformance criteria were not met (identified nonconformities)	shall	shall	shall	shall

Annex B
(informative)

Usability test report example

B.1 General

This is an example of an evaluation report.

Since the order in which the sections and the elements within an evaluation report are not prescribed, this example demonstrates that it is not necessary to follow the order in which the content elements were presented in this International Standard. However, the report does need to include all the required elements for conformance. This example evaluation report also illustrates that the grouping of the content elements themselves can be defined by the author of the report (e.g. combining information such as methods and procedures into one section of the evaluation report). All headings as introduced in this International Standard have been included, although they do not contain information in detail or in some cases include repetition of material. It is not necessary to use all the recommended headings in every evaluation report.

Evaluations often contain more than one type of evaluation as in this example evaluation (e.g. user observation and subsequent user survey). The content elements for each section of an evaluation report are determined by the type(s) of evaluation conducted. As a result this example evaluation report includes the content elements for both, observing user behaviour and user survey. The 'user survey' in this example uses the verbal comments made by participants as they completed each test as the basis for evaluation. The data, results and recommendations from both types of evaluation are presented together in the relevant sections of the report.

Within this example, the footnotes identify the conformance tables within this International Standard that apply to that section and demonstrate that the example report does include all the required elements for conformance.

This Annex has been adapted from an evaluation report provided by UXQB — The International Usability and User Experience Qualification Board. The website names and other identifying information (such as figures and screen shots) have been altered or eliminated for anonymity.

NOTE This example is not a complete report. It does not include all the findings or the figures which are mentioned within the example.

B.2 rentmytruck website usability test report

B.2.1 Executive summary[6)]

The running version of www.rentmytruck.com was usability tested in March 2011 with five members of the target group. The evaluation was based on observing user behaviour and the comments made by participants as they completed seven test tasks. Participants answered a set of questions about themselves at the end of the test. The test method was unmoderated remote usability testing, where typical users (test participants) carried out tasks while their actions and what they said were recorded. Screen and audio recordings of the test sessions were analysed by usability professionals after the test.

The purpose was to determine usability strengths and weaknesses of the rentmytruck website. This report describes findings and recommendations from the test.

6) Conforms to Table 3 for observing user behaviour and user survey.

Main positive usability findings for www.rentmytruck.com:

— **Back button always works without any problems**

Test participants always got what they expected when they pressed the Back button in their browser. On some comparable websites, the Back button does not always work as users expect.

— **The website preserves the contents of the shopping cart well**

Test participants never lost an item in the shopping cart. On some comparable websites, there is a timeout that clears the shopping cart and asks participants to start over. If a timeout limit exists on this website, none of our test participants encountered it.

Main improvement areas for www.rentmytruck.com:

— **Users can rent trucks without being asked about damage coverage**

Several participants rented a truck without ever seeing the Damage protection page. Their rental by default included no damage coverage, and the website did not ask them to make a choice.

— **Taxes, fees and total price are not shown**

Participants wanted to know the full price of the rental. The website only shows what users must pay in the store excluding taxes and fees. Participants were displeased that they could not see the total price in the shopping cart. They were even more displeased when they found out that they could not even see the total price on the checkout page.

B.3 Findings[7)]

B.3.1 General

B.3.1.1 Positive usability findings

— **Back button always works without any problems**

Test participants always got what they expected when they pressed the Back button. On some comparable websites the Back button does not always work as users expect.

— **The website preserves the contents of the shopping cart well**

Test participants never lost an item in the shopping cart. On some comparable websites, there is a timeout that clears the shopping cart and asks participants to start over. If a timeout limit exists on this website, none of our test participants encountered it.

B.3.1.2 Usability problem

— **FAQs use slang without explanation**

FAQs use slang without explanation: Examples of terms that confused test participants: Damage Waiver, Medical-life coverage, Liability coverage. User comments included:

"It would be nice if there was a little bit more to help me out with that"

"[The FAQs] didn't explain it well"

Recommendation

— Make terms that are hard to understand clickable. When users click these terms, show a pop-up that explains the meaning using commonly known terms and examples.

7) Conforms to Tables 13 and 14 for observing user behaviour and user survey.

B.3.2 Rent a Truck

B.3.2.1 Positive usability finding

— **Test participants easily realized how truck price depends on truck size**

See the example in Figure 1 (not included in this Annex).

B.3.2.2 Usability problem

— **Total price not shown on the checkout page**

Test participants were displeased that they could not see the total price in the shopping cart. They were even more displeased when they found out that they could not even see the total price on the checkout page. See Figure 3 (not included in this Annex). User comments included:

"'Due today' is $30.45. Where are they getting that from? It would be nice if they told me what this was like, what fees are being included here."

"Where is my final price?"

Recommendations:

— In the shopping cart and on the checkout page, show both what is due today and what the total price of the order including truck rental and storage rental.

— Offer prices both with and without taxes. If there is no room for both, omit the prices without taxes.

— Ask the user for the expected ending date of any storage rental included in the offer. Use the ending date to compute the total price of the storage rental.

— Make it completely clear how the website arrives at any total or subtotal. If there is insufficient room for a full explanation, add the link explain next to the total or subtotal. Clicking this link must display a detailed breakdown.

B.3.3 Insurance

B.3.3.1 Positive usability finding

— …

B.3.3.2 Usability problem

— …

B.4 Description of the object of evaluation[8)]

The website that was evaluated was www.rentmytruck.com, which was the one available to the public in March 2011. The home page of the website at the time of testing is shown in Figure 11 (not included in this Annex).

The target group for the website is individuals who are in possession of a valid driver's license and who want to rent a moving truck for a local or long distance move.

The website also sells and rents moving supplies and moving tools. It also provides storage space.

Test participants were explicitly asked to refrain from submitting orders. Apart from this restriction, the website was fully available to them.

8) Conforms to Table 4 for observing user behaviour and user survey.

The website is intended to be used in a typical office or home desktop environment.

B.5 Purpose of the evaluation[9)]

The purpose was to determine usability strengths and weaknesses of the rentmytruck website.

B.6 Evaluation method

B.6.1 Overview of the evaluation method[10)]

B.6.1.1 General

The usability evaluation focused on observing user behaviour. This usability test was conducted as an unmoderated remote usability test by the company <remote usability testing service>. This company specializes in unmoderated usability tests with users recruited from their user base. In an unmoderated usability test, users are not observed live while they carry out tasks. Instead, their interactions with the website and their verbal comments are audio recorded for later analysis.

Five users each carried out seven tasks on the website in separate test sessions. At the end of each test session, they answered a number of pre-defined questions about themselves.

Subsequently, the screen and audio recordings were analysed by three usability testers working independently. They each identified usability problems and strengths as well as the underlying usability defects. They then reviewed individual findings and agreed on a consolidated set of usability findings.

B.6.1.2 Methodological basis

This usability test used observation of the users' behaviour as evidenced by screen recordings. This enabled the identification of successful completion of tasks, together with problems encountered in carrying out the tasks. In addition, the recognized "think-aloud" method was used as a basis for the "user survey". This method is described for example in Dumas und Redish (1999): A Practical Guide to Usability Testing, and Hartson und Pyla (2012): The UX Book.

B.6.1.3 Test sessions

The evaluation is based on an analysis of screen recordings from five unmoderated test sessions of www.rentmytruck.com carried out in late March 2011 by <remote usability testing service>, together with analysis of audio recordings of users' comments.

Each test session lasted between 13 min and 23 min. The total time for each test session including answering questions was less than 30 min.

B.6.1.4 Target group for the system

The target group for the website is individuals who are in possession of a valid driver's license and who want to rent a moving truck for a local or long distance move. The target group includes a large part of the adult, English-speaking US population. The target group is expected to have some knowledge of computers and the internet, but they do not have to be computer professionals.

---

9) Conforms to Table 5 for observing user behaviour and user survey.

10) Conforms to Tables 6 and 7 for observing user behaviour and user survey.

Table B.1 — Usability test results for each test participant

Participant	Sex	Age	Occupation	Web savvy
1	M	24	Missionary	Average
2	M	52	Small business manager	Average
3	F	62	Retired. Formerly a television news producer, then licensed paralegal.	Average
4	F	36	Housewife	Average
5	M	31	Sales and marketing	Average

B.6.1.5 Recruitment of users

All test participants were recruited by the company <remote usability testing service>, which specializes in unattended usability testing. Both sexes were represented. Test participants were required to have a valid driver's license and to be considering renting a truck. Otherwise, no restrictions were imposed on the recruiting.

B.6.2 Usability test script[11)]

B.6.2.1 Briefing

<Remote usability testing service> does not publish the instructions given to test participants ahead of a usability test session.

B.6.2.2 General instructions given to the users

Each user was given instructions online for each of the seven tasks. They were asked to carry out the task and were told to verbally describe any problems or particular issues that they noticed.

B.6.2.3 Test tasks

The test tasks were defined to test the most frequently occurring tasks on the website based on website statistics. Tasks were reviewed by a user experience specialist working for rentmytruck and by several independent usability professionals. The usability of the test tasks and the instructions for the test participants in the unattended usability test were tested in two dry-runs. Test tasks and instructions were subsequently improved based on the feedback.

The following task set was used for all sessions. The tasks were carried out in the same order by each participant:

Scenario: Your friends Mike and Anna are about to move from Pittsburgh, PA to Denver, CO. They have an apartment in Pittsburgh consisting of a living room, a bedroom, a kitchen, and a bathroom. They want to find the cheapest service for the move to Colorado. They expect to make the move themselves with some help from a few friends.

They are planning to move out on April 14th and they expect the trip to take three days.

The couple plans to return to Pittsburgh after two years so they want to rent a self-storage unit in Pittsburgh for the stuff they do not need in Denver.

Task 1: The couple needs a truck that is suitable for all the furniture and belongings in their three-room apartment. Please find the total price the couple will have to pay for the truck.

NOTE They are moving on April 14th from XXX1 Rd. in Pittsburgh, PA 15217 to XXX2 St. in Denver, CO 80218.

11) Conforms to Tables 8 and 10 for observing user behaviour and user survey.

Expected answer: According to rentmytruck, a 14 ft truck is required. The price of the truck is $1,165 plus moving insurance $196 plus environmental fee $5. Taxes are not included. The tax rate does not seem to be available from the website.

Task 2: Before you go any further, you want to check if Mike and Anna need a special driver's license to drive the truck across country. Where would you find that info?

Expected answer: An ordinary driver's license is OK according to the FAQ "Do I need a special driver's license".

Task 3: They also need an indoor storage unit in Pittsburgh that can hold 10 moving boxes (18 in × 18 in × 16 in) and a large fridge. Find the per month cost of the storage.

Expected answer: The price of storage at XXX Self Storage, 1st floor, 5 ft × 5 ft × 8 ft for 24 months is $59 a month.

Task 4: […]

Expected answer: […]

Task 5: […]

Expected answer: […]

Task 6: […]

Expected answer: […]

Task 7: […]

Expected answer: […]

B.6.2.4 Task completion

Each task was considered complete as soon as each test participant moved on to the next task. (Since no moderator was present, it was not possible to help participants if they got stuck or when they arrived at an incorrect answer.)

B.6.2.5 Post-session questions

After each test session, the test participant answered several questions including the following:

— Are you male or female?

— How old are you?

— Have you ever rented a truck from this truck company before?

— Where do you live (town, state)?

— What is your occupation?

— How would you rate your web knowledge?

— Did you complete the tests at home or in your workplace, or other (if other, please specify)?

— Which type of device did you use, desktop computer or notebook/laptop (manufacturer and model)?

— Which browser did you use (including version)?

B.7 Findings for each test participant[12)]

For each test participant, the following table shows the analysis of the test participant's observed performance on each task.

Tasks were assessed and scored as follows:

— The task was solved correctly without problems; scored as 1.

— Problems occurred which delayed the test participant in carrying out the task; scored as 2.

— The test participant encountered considerable problems but eventually succeeded in completing the task correctly; scored as 3.

— The test participant was unable to complete the task or arrived at a result that deviated significantly from the correct result; scored as 4.

Table B.2 — Analysis of the test participant's observed performance

Participant	1	2	3	4	5
Task 1 — See notes. Price for truck rental, Pittsburgh to Denver	3	1	4	3	3
Task 2 Check need for special driver's license	1	1	1	1	1
Task 3 — See notes. Cost of indoor storage unit	1	3	3	3	3
Task 4 Phone number of nearest rentmytruck location	1	1	1	1	2
Task 5 — See notes. Rent truck and purchase moving supplies	4	4	4	4	4
Task 6 — See notes. Liability for repair costs	4	4	2	1	4
Task 7 Find nearest rentmytruck location	1	1	4	2	1

Notes to these findings

Problems in task 1: See the following findings in section 1.

— Insufficient help for selecting right truck size.

— Unclear if included mileage is sufficient.

— Taxes and fees are not shown.

— Total price not shown; website only shows what user must pay in store.

Problems in task 3: See the following finding in section 1.

— No help offered for selecting the right size of the storage room.

Problems in task 5: See the following findings in section 1.

— Adequate liability coverage is not included for all types of moves.

— Users can rent trucks without being asked about coverage.

12) Conforms to Table 12 for observing user behaviour and user survey.

Problems in task 6: See the following findings in section 1.

— Unclear if vandalism is covered by insurance coverage.

— Unclear if there is a deductible.

Bibliography

[1] ISO 9241 series, *Ergonomics of human-system interaction*

[2] ISO 26800:2011, *Ergonomics — General approach, principles and concepts*

[3] ISO/IEC/IEEE 15288:2015, *Systems and software engineering — System life cycle processes*

[4] ISO/IEC/IEEE 15289:2011[13)], *Systems and software engineering — Content of life-cycle information items (documentation)*

[5] ISO/IEC 17000:2004, *Conformity assessment — Vocabulary and general principles*

[6] ISO/IEC/IEEE 24765:2010, *Systems and software engineering — Vocabulary*

[7] ISO/IEC 25000:2014, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guide to SQuaRE*

[8] ISO/IEC 25010:2011, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*

[9] ISO/IEC 25040:2011, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Evaluation process*

[10] ISO/IEC 25063:2014, *Systems and software engineering — Systems and software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability: Context of use description*

[11] ISO/IEC 25064:2013, *Systems and software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability: User needs report*

[12] ISO/IEC 25065[14)], *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for Usability — User Requirements Specification*

[13] ISO/IEC 33020:2015, *Information technology — Process assessment — Process measurement framework for assessment of process capability*

[14] ISO TS 18152:2010, *Ergonomics of human-system interaction — Specification for the process assessment of human-system issues*

[15] ISO TR 18529:2000, *Ergonomics — Ergonomics of human-system interaction — Human-centred lifecycle process descriptions*

[16] ISO/IEC TR 11580:2007, *Information technology — Framework for describing user interface objects, actions and attributes*

[17] ISO/IEC TR 24774:2010*Systems and software engineering — Life cycle management — Guidelines for process description*

[18] ISO/IEC TR 25060:2010, *Systems and software engineering — Systems and software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability: General framework for usability-related information*

[19] IEC 62366–1:2015, *Medical devices — Part 1: Application of usability engineering to medical devices*

13) Withdrawn. Replaced with ISO/IEC/IEEE 15289:2015.

14) Under preparation.

ICS 35.080

Price based on 37 pages