VDE SPEC

Enhanced Glossary – Methods and Tools

VDE SPEC 90023 V1.0 (en)

VDE

FOREWORD

Publication date of this VDE SPEC: 07.02.2024

No draft has been published for the present VDE SPEC

This VDE SPEC resulted from the project "Harbsafe II" (funding code 03TN0018B) within the framework of the initiative WIPANO funded by the Federal Ministry for Economic Affairs and Climate Action (Bundesministerium für Wirtschaft und Klimaschutz, BMWK).

This VDE SPEC was developed according to the VDE SPEC procedure. VDE SPEC 90023 V1.0 (en) is developed in a project group and not necessarily with the involvement of all interested parties.

This VDE SPEC is not part of the VDE set of regulations or the German set of standards. In particular, this VDE SPEC is not a technical rule within the meaning of Section 49 EnWG.

The authors of this VDE SPEC are:

- Peter Bernard Ladkin, Causalis Ingenieurgesellschaft mbH
- Dieter Schnäpp, TU Braunschweig
- Mark Behrendt, Inosoft AG
- Sascha Man-Son Lee, DKE

Despite great efforts to ensure the correctness, reliability and precision of technical and non-technical descriptions, the VDE SPEC project group can neither explicitly nor implicitly guarantee the correctness of the document. This document is used in the knowledge that the VDE SPEC project group cannot be made liable for damage or loss of any kind. The application of the present VDE SPEC does not release the user from responsibility for their own actions and is therefore at their own risk.

In the course of the manufacture and / or introduction of products into the European internal market, the manufacturer shall carry out a risk analysis in order to first determine which risks the product may entail. After performing the risk analysis, he evaluates these risks and, if necessary, takes suitable measures to effectively eliminate or minimize the risks (risk assessment). The present VDE SPEC does not release the user from this responsibility.

Links to third-party websites do not constitute an approval of their content on the part of VDE. VDE is not responsible for the availability or the content of these websites. The establishment of a link to these websites is at the user's own risk.

Attention is drawn to the possibility that some elements of this document may affect patent rights. VDE, DKE, and DIN are not responsible for identifying any or all of the related patent rights.

Contents

Tables

1 Scope

1.1 General

This document describes methods supported by software tools for writing and maintaining technical terminology in ISO and IEC standards. The software tools are to be made available for use by standardisation committee members in the *Enhanced Glossary*.

The position and place of terminology in IEC standards is governed by IEC Directives (see Section 2: Normative References). Terminological definitions are normative (IEC Dir 2 16.2); mandatory (IEC Dir 2 16.3); and constitute Clause 3 of a standard (IEC Dir 2 16.4). The constitution and format of definitions shall be in accordance with ISO 10241-1:2011 (IEC Dir 2 16.5.2). Each <term, definition> subclause shall be numbered (IEC Dir 2 16.4).

In this document, the term definiens is also used for the definitional part of a terminological entry, as well as the term definiendum for the term being defined by the definiens. This is nominally contrary to the principle that only one term shall be used for a given concept. However, all four words, term, definition, definiens, and definiendum have common meanings and uses, and so according to IEC Directives Part 2, 16.5.5 would not normally be explicitly defined. However, they are central to terminology work and so it is appropriate to define them.

A major task in maintenance of technical terminology is harmonisation. Different branches of engineering, mathematics, and science use specific terms often in different ways. For example, the term cosine has a unique definition from elementary mathematics; whereas the term tangent has two: one for the ratio sine/cosine in trigonometry, and one in geometry for a straight line which is incident and parallel to a curve at a specific point on the curve. The term tangent is what is called a homonym, a term which has different meanings in different uses. In electrotechnology, many technical terms have different uses in different branches – the term risk is a notable example. In the IEC Glossary, a database of terminology from all IEC standards, there are many entries for risk. The International Electrotechnical Vocabulary (IEV) attempts to eliminate homonyms as far as possible.

It is neither possible nor desirable to ascribe one unique meaning to the mathematical term tangent, for the uses of this term are pervasively established in mathematics, science, and engineering for many hundreds of years. However, the term may be annotated by domain of use/field of use: tangent <trigonometry>; tangent <geometry> and these extended terms are then unique. Exactly the same technique is used in the IEV where necessary, and in principle throughout electrotechnical terminology. However, at time of publication, electrotechnical terminology is rife with homonyms (as may be seen by browsing the IEC Glossary). Harmonisation is desirable; principles are given in ISO 860:2007.

1.2 Content

1.2.1 General

Three methods are described in this document to aid terminology work and harmonisation: SemAn, HintAn, and STS. SemAn is supported by the tool SemAn Analyser. HintAn is supported by the tool HintAn Analyser. STS is a pure tool. They are described in detail in Section 5.

1.2.2 SemAn and the SemAn Analyser

SemAn is a method for semantic analysis of the definiens. It attempts to translate the definiens into a form isomorphic to a term in the language of sorted first-order logic. SemAn is described in 5.2.

SemAn is supported by a software tool, the SemAn Analyser, which takes the natural-language definiens, annotates with the logical constants [AND] and [OR] where appropriate, and pretty-prints/aligns the output. It treats all words and punctuation occurring in the definiens as primitive. It does not attempt to sub-analyse these further, as an application of SemAn might. The SemAn Analyser is described by example in 5.3.

1.2.3 HintAn and the HintAn Analyser

IEC Dir 2 and ISO 10241-1:2011 contain requirements for the form and formatting of the definiens. HintAn checks 19 of these requirements. These are listed and briefly explained in 5.4. Section 5.5 contains examples of their use in IEC terminology. HintAn may be used as a checklist for and during the formulation of definiens by a standardisation committee to ensure conformance with IEC Dir 2 and ISO 10241-1:2011.

The HintAn Analyser is a software tool which checks definitions according to HintAn. Not all HintAn criteria are implemented in the HintAn Analyser; but many are.

1.2.4 STS and Structural Similarity

It is possible by means of SemAn to analyse the definiens of various homonyms, to thereby identify small, moderate, or large differences and to classify the various definiens by similarity. For example,

- some terms are variously defined in different documents with differences that amount to minor syntactic adjustments and may easily be brought together into one syntactic form,
- some terms exhibit minor semantic variations which could in principle be brought together via straightforward committee negotiations; for example, harm and asset, both discussed in detail using SemAn in 5.2,
- some terms exhibit substantial differences in definiens. For each, either a new definiendum has to be chosen for one, or the term must be qualified by domain of use. This task can become problematic because of the strong cultural attachments to specific terms in widespread use in a single branch.

STS is a software tool based on machine-learning software which estimates definiens similarity according to these three broad categories, and

- attempts to collect terms with syntactic or minor semantic differences (as indicated by similar overall syntax with varying elements) together into a semantic class,
- exhibits the terms grouped into their semantic classes.

STS thus performs part of the task on homonyms with many variants that otherwise would have to be performed manually by SemAn/SemAn Analyser. In this way, STS specifies the harmonisation task for a homonym. STS does not, however, attempt to solve this task; the task is solved by SemAn, or the use of ConcAn (conceptual analysis).

Since STS is a pure software tool, there is no manual method associated with it. This document therefore does not include a separate section describing any method behind STS.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 60050, *International Electrotechnical Vocabulary (IEV). Many sections, many dates. Online-Version: http://www.electropedia.org/*

IEC Glossary (no date), *Online at http://std.iec.ch/glossary*

ISO 10241-1:2011, *Terminological entries in standards – Part 1: General requirements and examples of presentation. First edition, 2011-04-15, confirmed 2022*

ISO 10241-2:2012, *Terminological entries in standards – Part 2: Adoption of standardized terminological entries. First edition, 2012-08-15*

ISO 860:2007, *Terminology work – Harmonization of concepts and terms. Edition 3, 2007-11 confirmed 2021*

ISO/IEC Directives Part 1 + IEC Supplement, Edition 17.0: 2021-05, *Procedures for the technical work – Procedures specific to IEC*

ISO/IEC Directives Part 2, Edition 9.0: 2021-05, *Principles and rules for the structure and drafting of ISO and IEC documents*

3 Terms and definitions

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

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

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

– IEC Electropedia: available at https://www.electropedia.org/

3.1
alphabet <Indo-European languages>
set of symbols

3.2
definiens
definition
phrase specifying the meaning of a definiendum which can be substituted salva veritate for that term in any sentence in a standard

3.3
definition
definiens
phrase defining a technical term

Note 1 to Entry: A definiens must usually satisfy the substitution requirement and other requirements in ISO/IEC Directives Part 2: 2021, 16.5.

3.4
definiendum
term
word or short phrase with a technical meaning in the field of use

3.5
domain of us
field of use
area or branch of engineering, science, mathematics, or other intellectual discipline in which there exist branch-specific meanings of terms

3.6
export format
format for presentation of data to users of an on-line service

3.7
first-order logic (FOL)
predicate logic
formal logic of objects, predicates, and quantifiers using the logic constants AND, OR, NOT, IMPLIES, SOME, ALL

Note 1 to Entry: There are many formal systems which encode FOL. Only its underlying language is used in SemAn.

3.8
formal language
set of strings of symbols, with a rigorous formation rule or rules which allow effective determination of which strings belong to the language and which do not

3.9
harmonisation <concept>
activity leading to the establishment of a correspondence between two or more closely related or overlapping concepts having professional, technical, scientific, social, economic, linguistic, cultural, or other differences, in order to eliminate or reduce minor differences between them

[SOURCE: ISO 860:2007, 3.1]

3.10
harmonisation <definition>
activity leading to the description of a harmonized concept by an intentional definition that reflects the position of the concept in the harmonized concept system

[SOURCE: ISO 860:2007, 3.3]

3.11
harmonisation <term>
activity leading to the selection of designations for a harmonized concept either in different languages or within the same language

[SOURCE: ISO 860:2007, 3.4]

3.12
homonym
term with multiple non-identical definiens in multiple standards

3.13
logical constant <FOL>
AND, OR, NOT or ∀ (“for all”) or ∃ (“there exists”)

3.14
machine learning engine (MLE)
software which accomplishes a task by using machine-learning techniques, having been trained on example input

3.15
meaning postulate (MP)
assertion, assumption, or hypothesis used in reducing and reworking a putative definiens into a simpler or more preferable form

3.16
primitive
unanalysed term or phrase

3.17
semantic analysis <IEC terminology>
analysis according to the principles of SemAn

3.18
semantic class
collection of differing definiens associated with a specific homonym that exhibit either minor syntactic variation or minor semantic variation that is judged to be relatively easy to reconcile

3.19
sentence <formal logic>
member of the formal language

Note 1 to Entry: A sentence will usually be assigned a truth value (TRUE or FALSE) by the formal semantics of the language. The truth value is usually determined by a particular interpretation, but some sentences usually have constant truth values (namely the logical truths and logical falsehoods).

3.20
sentence <Indo-European natural language>
string of symbols used to make an assertion, a query, or other speech act

3.21
sorted first-order logic
sorted FOL
FOL with a distinguished finite set of unary predicate symbols called sorts, along with the axiom that any object has a sort

Note 1 to Entry: If the sorts are S1, …., Sn, then the axiom is ∀x(S1(x) OR S2(x) OR … OR Sk(x)).

Note 2 to Entry: The sorts represent the common types of object: “a man” is in sorted FOL “a thing which is a man”; “a component of system S” is “a thing which is a component of system S”.

3.22
substitution requirement
requirement that a definiens be able to substitute syntactically for its definiendum in any sentence without change of meaning of the sentence

Note 1 to Entry: This requirement is established by ISO/IEC Directives Part 2: 2021, 16.5.6.

3.23
symbol <formal language>
formal symbol
primitive element of a formal language

3.24
symbol <Indo-European natural language>
member of the alphabet of the language

Note 1 to Entry: IEC standards are published in English (and sometimes French). Indo-European languages are constructed as strings of symbols, albeit from somewhat different alphabets.

3.25
syntactic unit <Teutonic or Romance languages>
item of syntax <Teutonic or Romance languages>
a single word, alternatively a number of associated words, such as article+noun, determiner+noun, adverb+verb, which express a concept

Examples to Entry: “a switch”, “the middle point”, “the lowest positive value”, “combination”, “slowly rising”, “continuously rotating wheel”

3.26
word <Indo-European natural language>
string of consecutive symbols without spaces, forming a linguistic unit

4 Symbols and Abbreviations

Abbreviation	Signification
FOL	First-order logic
HintAn	Hint Analysis; method for apprising deviations from the IEC Directives and ISO 10241-1:2011 in the format and presentation of definitions. The HintAn Analyser is a demonstration tool, which implements most of HintAn.
IEV	International Electrotechnical Vocabulary (IEC 60050, many parts)
LFOL	Language of First Order Logic
MLE:	Machine-learning engine
MP:	Meaning postulate
SemAn:	Semantic Analysis; method for the semantic analysis of definitions. The SemAn Analyser is the Tool for annotation/pretty printing of definitions according to SemAn.
STS:	Semantic Textual Similarity; tool based on MLE for grouping varying definitions of a homonym into semantic classes.
TC:	Technical Committee

5 The Enhanced Glossary, SemAn, HintAn and STS

5.1 The Enhanced Glossary Annotated Properties

5.1.1 Heading

The first item/function is the **heading**, which states;

1. the norm the selected definition appears in,

2. which committee accounts for this definition,
3. its publication ID.

The heading is followed by the **term** (definiendum) with synonyms, the **definition** itself (definiens) and the **notes** to the entry.

This information is identical to that given in the IEC Glossary.

Then follow additional features of the Enhanced Glossary. These are illustrated through the example of the term signal from Clause 3 of IEC 61671:2012.

5.1.2 Running Example: signal

In this section, a running example is used to illustrate functions available in the DKE Enhanced Glossary. The *heading* is:

> **Signal**
>
> (A) An electrical impulse controlled or observed by a test resource. (B) A visual, audible, or other indication used to convey information.

5.1.3 Applying HintAn

It is well to check first whether a definiens conforms with the required format of a definition according to IEC Directives Part 2 Section 16 and ISO 10241-1:2011. The definiens of *signal* is not so conformant.

The method HintAn gives information, derived from ISO/IEC Directives on terminology as well as the standard ISO 10241-1:2011 on the wording of the definition at hand and what would improve conformance with the Directives and named standard. The HintAn function is implemented in the DKE Enhanced Glossary by the software tool HintAn Analyser.

The end decision on what is and is not appropriate ultimately stays with the working group/maintenance team of the responsible IEC Technical Committee (TC), possibly in consultation with TC 1 (the Terminology TC). In the case of the running example, the responsible TC is TC91.

For *signal,* the HintAn Analyser produces the following *Hints*:

- [Determiner],
- [Too Many Definitions],
- [Sentence or Multiple Sentences],
- [Definition Ends with a Full Stop or Starts with Caps],
- [Definitions Vary a Lot].

This translates as follows.

- According to ISO/IEC Directive 2 a definiens should not start with an article: the “An” at the start of the definition text is deprecated (ISO/IEC Directives Part 2:2021, 16.5.6.)
- The definiens contains two separate definitions (visibly marked by the enumeration). According to ISO/IEC Directive 2, a definiens should be solitary and not multiple (also ISO/IEC Directives Part 2:2021, 16.5.6; implied by the *substitution requirement*) To resolve this doubling and decide which characteristics are needed to define the concept behind the term, expert domain knowledge is here required.
- The third advice, that the definiens contains sentences (and thus vitiates the substitution requirement) will be resolved when the multiple definitions are resolved into separate definitions.
- The fourth advice (also derived from the substitution requirement) is resolved by commencing the definiens in lower case and omitting the full stop termination at the end.
- Finally, a relational *Hint* informs about the fact that the definitions for this term “vary a lot”, meaning that a automatic similarity check has concluded there is significant divergence. This can be read, on the one hand, as an incentive to revise the presentation (which is already present in the previous hints) and, on the other hand, as a caveat to be especially careful with these entries: a multitude of diverging definitions could be a sign of the term’s concept not being as clear as usually required.

5.1.4 Manually Correcting the Example

Following the hints raised by HintAn leads straightforwardly to the following revised definition (pair of definitions).

Signal

<electrical> electrical impulse controlled or observed by a test resource.

Signal

<information> visual, audible, or other indication used to convey information.

These definitions are conformant, and the definiens are identical to the wording of the original. A specialist may, however, query whether the suggested field of use is adequate, and may propose another (pair).

5.1.5 Identical Definitions

The third function in the *DKE Enhanced Glossary* produces a list of all the other norms the term appears in together with the same definition the user has selected for export. For *signal* there is no further co-occurence with the identical definiens.

5.1.6 Homonyms

The fourth special feature “Definition inconsistencies”, provides detail on additional appearances of the definiendum whose definiens do not coincide with the one being analysed: that is, homonyms. Homonyms are explicitly denigrated in the IEC Directives and ISO 10241-1 and a resolution is required. The presentation of the varied definitions associated with the homonym term allows for an effective overview, as follows.

The findings are divided up into semantic classes by similarity of definiens. This is achieved through a software engine STS developed using machine learning (i.e., an MLE). To determine whether two definitions belong in the same or different semantic classes, STS classifies a given pair of definitions according to the following scheme:

1. there is a significant semantic difference;
2. there is a less significant, but still perceptible, difference (e.g. when one definition entails the other);
3. there is a general phrase equivalence (minor semantic difference);
4. the variation is syntactic (or largely syntactic).

Comparisons resulting in 1 or 2 result in the two definiens being assigned to different semantic classes. Comparisons resulting in 3 or 4 result in the two definiens being assigned to the same semantic class. An example classification of 18 varying definitions for the term “*risk*” are given in 5.6, Table 26.

For some terms, there is a vast amount of “competing” definitions, making it difficult to survey in its entirety. The use of semantic classes and automatic classification via STS renders the informational load manageable for the human eye and mind, as may be seen further in 5.6.

5.1.7 Analysing Definiens with the Help of SemAn

The final function provides a semantic analysis of a definiens, using the SemAn Analyser, presenting the output as an annotated, pretty-printed version of the definiens, including all punctuation. All words in the definiens are considered as primitives by the SemAn Analyser. The method SemAn itself, described in 5.2, may render a word into its semantic components, for example if the word is a defined term elsewhere in the glossary, but the SemAn Analyser does not currently do that; there is a manual step involved. After manual replacement of such a defined term, the SemAn Analyser could be re-invoked on the new definiens.

HintAn Analyser has suggested issues with the original definition of signal, and modifications conformant with the terminology directives and standards have been proposed. The result of invoked the SemAn Analyser on the two proposed conformant definitions of signal is

```
signal:
<electrical> electrical impulse controlled or observed by a test resource
\\
```

```
electrical impulse
                  > [OR] controlled
                  > [OR] or observed
                           > by a test resource

signal:
<information> visual, audible, or other indication used to convey
information
\\
     > [OR] visual
     > [OR] , audible
     > [OR] , or other indication
                        > used to convey information.
```

SemAn Analyser exhibits the first as a sort, *electrical impulse*, with one of two relations to a *test resource*, that of being *controlled* or that of being *observed*. The second definition is a disjunction: the qualifier of all three disjuncts is that they are *used to convey information*, and the information conveyer may be *visual* or *audible* or an *other indication*. In this case, with simple and short definitions which have no key terms in common, a comparison of the two definiens shows that they are clearly distinct concepts.

SemAn and the SemAn Analyser output are further considered in 5.2 and 5.3 respectively.

This concludes the brief synopsis of the analytical and informational functions available through the DKE Enhanced Glossary, with the running example of *signal*. The methods SemAn and HintAn, and the software implementations SemAn Analyser, HintAn Analyser, and STS are considered in more detail below.

5.2 Semantic Analysis by Means of SemAn

5.2.1 General

Formal semantic analysis is a well-developed subject in linguistics, nowadays often a branch of applied formal logic. The specific kind of semantic analysis described here is given the name SemAn and supported by a software tool, the SemAn Analyser.

5.2.2 Semantic Analysis: Preliminaries

The term “semantic analysis” is used here as a technical term which refers to a specific way in which definitions in technical terminology may be analysed. The SemAn method has been developed specifically for electrotechnical terminology occurring in Clause 3, “Terms and Definitions” of IEC standards. Manual examples of SemAn are given first, below, to show the method. Output of the SemAn Analyser is formatted (“pretty-printed”) text with annotations illustrating logical structure.

SemAn is particularly geared towards comparative analysis, in which one has syntactically-varying definitions of the same term (homonyms) or syntactically-similar definitions of different terms (quasi-synonyms). SemAn allows the similarities and divergences between the terms to be illustrated in a canonical and intuitive way. SemAn exhibits the logico-semantic structure of individual natural-language definitions, so it also enables individual definitions to be improved to enhance understanding.

There is no one unique resulting analysis of a definition in SemAn. One may choose different primitives (unanalysed words or phrases): in one analysis, a syntactic unit may be taken to be primitive; in another analysis, a slight divergence of that syntactic unit from another item in a related definition may require the unit be further analysed (as a compound of further primitives) so that the divergence can be exactly specified.

The software tool SemAn Analyser gives one output per conformant definition, illustrating the logical form of the definition while taking the individual syntactic units to be primitive. It annotates the definiens with logical constants and punctuation and pretty prints it, as for example in 5.2.5. Further examples of SemAn Analyser output are given in 5.3.

5.2.3 The Formal Language of SemAn

Formal semantic analysis in the linguistics of natural language, as it is practiced today, uses formal annotation into which a target definition is parsed. So does SemAn. The language used is isomorphic to the language of first-order logic (FOL). An introduction to the language of propositional logic and

FOL is to be found in [1]. Nearly a century and a half of experience with FOL has established its preeminence as a system in which assertions may be made with precision and formal inferences may be precisely codified (there are other logics, often known as higher-order or non-classical logics depending on their type, which are useful for similar purposes in domains in which FOL is limited). The language of FOL (LFOL) consists of

- predicate symbols,
- object symbols (divided into constants, which SemAn uses, and variables, which SemAn does not use),
- functional symbols (largely not used in SemAn),
- the logical constants AND, OR (used widely in SemAn and SemAn Analyser),
- the logical constant NOT (largely not used in SemAn and SemAn Analyser, because negations are often incorporated into the terms themselves),
- quantifiers (like negation, largely incorporated into the syntactic items themselves).

The meaningful syntactic units of LFOL are sentences. There are no meaningful parts of sentences such as phrases which are not themselves sentences. This entails that translating natural language expressions into LFOL requires expanding the expressions to conform with the phraseology of LFOL.

5.2.4 Translating natural language phrases into the language of SemAn

As far as is yet known, there is no generally accepted algorithm for translating natural language sentences into LFOL in a way which preserves their meaning. However, there are some more or less standard translation rules, partly illustrated below.

> (English) *John or Joan opened the front door.*

First, the phrase in subject position, *John or Joan,* has no equivalent in LFOL. In LFOL, *OR* may only be used to conjoin sentences. Second, there are no syntactic elements corresponding to phrases in LFOL, only sentences and their component symbols. Third, the sentence intuitively speaks to one of two situations; one in which John opened the front door, and another in which Joan opened the front door (as well as a third in which they both did, but presumably not simultaneously). These observations may be used to convert the English sentence into one conforming to LFOL with the same intuitive meaning, namely:

> (LFOL) *John opened the front door OR Joan opened the front door.*

In English, phrases in subject position or object position can also be lists, with one constant (usually separating the last two list words) and separated by commas, as in:

> (English) *John, Joan or Jeremiah opened the front door.*

Similar principles apply here as above, and we obtain the translation:

> (LFOL) *John opened the front door OR Joan opened the front door OR Jeremiah opened the front door.*

A further step is that of constructing synonyms for predicates. In this example, three different people seem to be engaging in the same action, "*opened the front door*". In LFOL, the following action can be performed. A simpler symbol may be used to stand for the verb phrase "*opened the front door*". Second, whereas in English the subject (the person who opened the door) is typically written first and the predicate (the action) follows, with no punctuation, as in *John opened the front door*, in LFOL the assertion is expressed in a symbolic form akin to that of the elementary mathematics of functions: the argument (whoever did the opening) is expressed in parentheses after the predicate, as in *opened-the-front-door(John)* (here, hyphens are used to indicate that the predicate is denoted by a string of words rather than a single word). When symbol P is chosen to represent *opened the front door*, then this becomes syntactically easier to read.

> (LFOL) Let the symbol *P* stand for the predicate "*opened the front door*". Then the assertion becomes.
>
> *P(John) OR P(Joan) OR P(Jeremiah).*

SemAn uses such a natural-language version of LFOL as has been illustrated above. By experience, the illustrated translations seem to cover the routine majority of the task of translation. This language will become more clear when examples are discussed below.

5.2.5 The SemAn Analyser

The SemAn Analyser, in contrast to (manual) SemAn, does not use LFOL at all. It parses, annotates and pretty-prints the words and punctuation in the definition itself, in the order in which they occur. A manual translation from SemAn Analyser output into LFOL is intended to be straightforward. If a translation is not straightforward, this serves as an indication that the original definition proposed may be deficient, say unclear. The recommended remedy is to modify the source definition so that the translation of SemAn Analyser output into LFOL becomes straightforward.

The SemAn Analyser:

- Takes all English words as primitive formal symbols.
- Exhibits the logical structure of phrases by means of annotations using the logical constants (AND, OR, more rarely NOT) and marked indents.

For example, the annotated/pretty-printed output of the SemAn Analyser invoked on the term *harm* is as follows:

```
67.
harm:
physical injury or damage to the health of people or damage to property or
the environment
\\
     [OR] physical injury
     [OR] or damage
               > [OR] to the health
                           > of people
               > [OR] to property
               > [OR] or the environment
[Source: IEC 61508-4:2010].
```

It can be seen that the SemAn Analyser takes the definiens as syntactically given and marks it up. It treats *physical injury* as a primitive. Below, a manual SemAn does not take this phrase as primitive and compares with *damage to the health of people* in order to determine if the definition can be expressed more succinctly and clearly.

It follows that the output of the SemAn Analyser is not a full SemAn, but a preliminary processing of the definition which exhibits certain formal features of the definition, enabling improvements to be made where they are appropriate, and which enables a human analyst to continue the SemAn if desired, for example by considering the meaning of *physical injury* and relating it to *damage to the health of people*.

5.3 Examples of SemAn Analyser Output and of SemAn

5.3.1 Output of SemAn Analyser on *asset*

There are two non-identical definitions of asset in the IEC 62443 series of standards. Both are considered below, in order to illustrate the harmonisation task, and to show how much easier it is made by using the SemAn Analyser.

SemAn Analyser output on the two definitions of *asset* is:

```
10.
asset:
physical or logical object owned by or under the custodial duties of an
organization, having either a perceived or actual value to the organization
\\
     physical or logical object
          > [AND] owned by or under the custodial duties
                    > of an organization
          > [AND] having either a perceived or actual value
                    > to the organization
[Source: IEC 62443-2-1:2010]
[Source: IEC TS 62443-1-1:2009]
11.
asset:
physical or logical object having either a perceived or actual value to the IACS
\\
     physical or logical object
```

```
          > having either a perceived or actual value
                > to the IACS
[Source: IEC 62443-3-3:2013].
```

This annotated parsing/pretty-printing immediately shows a number of similarities and differences in the two definitions. First, an *asset* is a *physical or logical object*. Second, it *[has] a perceived or actual value.* To whom the value accrues is different in the two cases (some implicit *organisation* in the first, presumably a human organisation such as a company; in the second, a system, namely the *IACS*). Similarly, the first definition mentions custodial duties associated with the asset; the second mentions no such duties.

This comparison gives clear indications of difference, and therefore the scope of discussion, to domain experts attempted to harmonise the two definitions. The harmonisation task here is twofold.

- To whom/what does the value of the *asset* accrue?
- Is the ownership/custody of the *asset* a key property? Is is implicit, or does it need to be explicit?

5.3.2 Example: a Manual SemAn of *harm*

1. Definition: harm IEC 61508-4 subclause 3.2.1 and IEC Guide 120 subclause 3.7:

   physical injury or damage to the health of people or damage to property or the environment.

SemAn goes further than the SemAn Analyser, using domain knowledge about the concepts (words and phrases) occurring in the definition (recall that the SemAn Analyser takes these as primitive). Invoking domain knowledge results in a meaning postulate. Because the result analysis has used the meaning postulates, they are restated along with the result of the SemAn.

First fill this definition out by expanding conflations (see 5.3.3).

- Expand syntactic conflation: “OR” is used to conjoin two noun phrases. The SemAn Analyser has identified two “levels” of conjoined phrase:
  - associated with *physical injury,*
  - associated with *damage.*

A first step is thus to expand. The physical injury or the damage is associated with the same qualifying phrase, namely:

to the health of people OR to property OR to the environment.

There are two ways this qualifying phrase can be treated:

- Parentheses can be used to make a unit out of this phrase;

  (*to the health of people OR to property OR to the environment*).
- An auxiliary definition can be used:

  Let P stand for *to the health of people OR to property OR to the environment.*

- The resulting phrase is

  physical injury (to the health of people OR to property OR to the environment)

  OR

  damage (to the health of people OR to property OR to the environment).

  Alternatively

  P(physical injury) OR P(damage).

The second alternative is obviously of no help whatever in further analysis. The first alternative is used to proceed.

- Consider next the first conjunct:

  physical injury (to the health of people OR to property OR to the environment).

  The ORs can be expanded further:

  physical injury to the health of people

OR

physical injury to property

OR

physical injury to the environment.

Semantic domain knowledge is invoked: (meaning postulate MP1) only people or sentient beings can be *physically injured*, not property or the environment. According to this, this may be further reduced:

physical injury to the health of people.

Using further domain knowledge, we note that *physical injury to the health* is redundant:

physical injury to people.

- Consider the second conjunct:

  damage (to the health of people OR to property OR to the environment).

  Again, this expands to:

  damage to the health of people

  OR

  damage to property

  OR

  damage to the environment.

- Conjoining the two expanded/reduced phrases gives

  physical injury to people

  OR

  (damage to the health of people

  OR

  damage to property

  OR

  damage to the environment).

  Note that logical *OR* is associative: *A OR (B OR C)* is the same as *(A OR B) OR C* and thus either may be written unambiguously without parentheses: *A OR B OR C* [1]. So this can be written:

  physical injury to people

  OR

  damage to the health of people

  OR

  damage to property

  OR

  damage to the environment.

- *Physical injury* is a term which contains *injury*, and (meaning postulate MP2, obviously related to MP1) injury can only occur to sentient beings. The term *people* is used; (domain knowledge) *people* is a plural of *person*, as is *persons*. The question arises if harm can be caused to one person, or must it always be more than one (plural)? Singular or plural? (Meaning postulate from domain knowledge MP3) Harm to one person is still harm. The issue could be clarified by rewriting *people as one or more persons:*

  physical injury to one or more persons

  OR

damage to the health of one or more persons

OR

damage to property

OR

damage to the environment.

- The first two conjoined clauses have as part *one or more persons*. Furthermore, they are semantically related: (MP4) *Physical injury is damage to the health* of (a person or persons). But is all *damage to the health* also *physical injury*? No, there can be damage to health that is predominantly psychiatric: post-traumatic stress syndrome for example. So (MP5) damage to the health includes physical injury but not vice versa. Put in terms of logic,

  physical injury to one or more persons

  IMPLIES

  damage to the health of one or more persons

  but not vice versa. It follows that the first clause can be omitted without semantic loss. However, an analyst might wish to retain it as a means of emphasis.

- Result:

  damage to the health of one or more persons

  OR

  damage to property

  OR

  damage to the environment.

  Alternatively:

  physical injury to one or more persons

  OR

  other damage to the health of one or more persons

  OR

  damage to property

  OR

  damage to the environment.

- Finally, these could be consolidated, by regrouping according to English conventions, for example

  physical injury or other damage to the health of one or more persons

  OR

  damage to property or to the environment.

  Alternatively:

  damage to the health of one or more persons, or to property, or to the environment.

The second definition can now be considered.

2. Definition: harm IEC Guide 51 subclause 3.1:

   injury or damage to the health of people, or damage to property or the environment.

- It is clear from the preceding analysis that;
  - the analysis can proceed largely as before,
  - (MP6) *damage to the health* can be considered equivalent to *injury.*
- Result:

damage to the health of people, or damage to property or the environment.

Equivalently:

injury to people, or damage to property or the environment.

Given that *people* and *one or more persons* are synonyms, as are *injury to* and *damage to the health of*, it follows that, under MP1, …, MP6, the two definitions are synonymous. The results may be expressed as follows:

- Under meaning postulates;
  - (MP1) only people or sentient beings can be *physically injured*, not property or the environment,
  - (MP2) injury can only occur to sentient beings,
  - (MP3) harm to one person is still harm,
  - (MP4) *Physical injury* is *damage to the health* of (a person or persons),
  - (MP5) damage to the health includes physical injury but not vice versa,
  - (MP6) *damage to the health* can be considered equivalent to *injury,* the two definitions are synonymous and equivalent to *physical injury or other damage to the health of one or more persons, or damage to property or to the environment damage to the health of one or more persons, or to property, or to the environment injury to people, or damage to property or the environment.*
- It follows that there is a harmonisation task, but one which is in this case purely syntactic: an analyst must choose between the three example definitions above (or ones in which synonymic phrases are used, such as *people* instead of *one or more persons* or vice versa).

This analysis is laborious, and the result relatively easily foreseeable from the start, but the purpose is to illustrate the principles and steps involved in a manual SemAn, including the formulation of MPs and this is shown more easily on such straightforward examples.

5.3.3 SemAn Example: *asset*

1. asset, IEC 62443-1-1 subclause 3.2.6 and IEC 62443-2-1 subclause 3.1.3;

   physical or logical object owned by or under the custodial duties of an organization, having either a perceived or actual value to the organization.

The SemAn proceeds with similar steps to that of harm in 5.4.3. There are not elaborated here in as much detail. However, the SemAn itself is more complex.

- Expand: An *asset* is a *physical object* OR a *logical object* <*with additional properties*>. The adjectives here are (meaning postulate MP1) applicative, so an *asset* is an object. It is left unexplained exactly what a logical object is. (One might speculate that a *metaphysical object* is meant, but engineers do not use that term.) Introduce the primitive *Ob* to denote the thing of sort *Obj* which is being talked about. The mathematics-type notation typical of formal logic is used: *P(Ob)* says *Ob is physical*, *L(Ob)* says *Ob is logical*. For objecthood, then, the term *Ob* of sort Obj has been introduced and yields the assertion *P(Ob) OR L(Ob)*.
- Fill out <*additional properties*>: another OR syntactic conflation is expanded. *owned by an organisation* OR *under the custodial duties of an organisation*. There is another sort here, *organisation*. Whereas for *Ob*, a specific asset is meant, the organisation is unspecified: (MP2) *some organisation* is meant. The term *some* is a quantifier and in logic, one would be tempted to quantify: "*there is an organisation Org such that ….*". But for a given asset, it can be assumed that (meaning postulate MP3) there is just one organisation that owns it or just one organisation that has custody of it. Note that this meaning postulate is not like the ones involved in the analysis of *harm*; it involves rather an assumption about the way of the world; that if there are multiple owners or custodians, just one can be singled out to be *Org* for the purposes of the definition. So, for *Ob*, there is a single *Org* of sort *Organisation* (let us say *Orgn*) which either *Owns* it or *HasCustody* of it: *Owns(Ob,Org) OR HasCustody(Ob,Org)*.
- Fill out ",": there is a list of properties here, starting with *owned by… OR under the custodial duties of …* and then *having ….* It is clear that *AND* is meant by the comma.

- Fill out further. Result: *(Ob has a perceived value to Org) OR (Ob has an actual value to Org)*. Choose primitives PV and AV for the predicates *has a perceived value to* and *has an actual value to*. The result is *PV(Ob,Org) OR AV(Ob,Org).*
- Result: it seems the analysis has arrived at the following, in formal form:

  P(Ob) OR L(Ob) AND Owns(Ob,Org) OR HasCustody(Ob,Org) AND PV(Ob,Org) OR AV (Ob,Org).

  However, there is an AND/OR ambiguity which needs to be disambiguated. AND/OR ambiguities arise because *A AND (B OR C)* does not have the same meaning as *(A AND B) OR C* and when there are no parentheses, as in *A AND B OR C*, one cannot tell which is meant. To disambiguate, parentheses are used:

  (P(Ob) OR L(Ob)) AND (Owns(Ob,Org) OR HasCustody(Ob,Org)) AND (PV(Ob,Org) OR AV (Ob,Org)).
- Rewriting the result:

  This formula looks “formal” and is typical for the indication of the logical structure of phrases and sentences/assertions. But it is hard to read. There are some ways to make such formulas easier to read, for example the vertical stacking of clauses, as in TLA [2]. In the TLA “pretty-printing” style, all clauses in a conjunction are preceded by the conjunction sign and stacked vertically, mutatis mutandis for disjunction. Indentation allows the elimination of the parentheses used for disambiguation:

  && P(Ob) OR L(Ob),

  && Owns(Ob,Org) OR HasCustody(Ob,Org),

  && PV(Ob,Org) OR AV (Ob,Org).

  The OR clauses within the conjuncts can be similarly formatted if so wished (the symbol ∨ is used to denote OR), to yield:

  && ∨ P(Ob),

  ∨ L(Ob),

  && ∨ Owns(Ob,Org,

  ∨ HasCustody(Ob,Org),

  && ∨ PV(Ob,Org),

  ∨ AV (Ob,Org),

  but there seems to be little point to doing so here. It is up to the analyst to decide which is most helpful. The sorts of *Ob* and *Org* have been so far left implicit, but there might be circumstances in which one needs to reason with them taken into account (see below). When introduced, the formal sentence in the language of sorted logic looks like:

  && Obj(Ob),

  && Orgn(Org),

  && P(Ob) OR L(Ob),

  && Owns(Ob,Org) OR HasCustody(Ob,Org),

  && PV(Ob,Org) OR AV(Ob,Org).

Consider now the second definition of asset.

2. asset IEC 62443-3-3 subclause 3.1.1

   physical or logical object having either a perceived or actual value to the IACS.

- Fill it out: The first observation is that too much was done with the first definition. The predicate *physical* need not have been separated from the predicate *logical*: instead of *P(Ob) OR L(Ob)* we could have used one predicate *PorL(Ob)*. But no matter; it was done and will be left so.
- Fill it out: again, *perceived value OR actual value*, but the subject of the valuation has changed. Now, it is not an *organisation* (a group of people) but is an engineering object, a system, namely the Industrial Automation and Control System – IACS - to which the IEC

62443 series is specifically targeted. Here, (meaning postulate MP3) there is no possible ambiguity as to which IACS is meant: it is the one to which this standard is currently being applied. A sort *IACS* is introduced along with a primitive *theIACS* for an object of this sort.

- Result:

  && Obj(Ob),

  && IACS(theIACS),

  && P(Ob) OR L(Ob,

  && PV(Ob,theIACS) OR AV(Ob,theIACS).

There are now two analysed definitions of *asset*, which are not identical. The term *asset* is thus a homonym. The task of harmonisation is to select one of these as the primary definition. There are most often two ways in which this may be done. First, definitions may be specialised to domains of application, as explained in 5.1. So, for example, s*ignal* means one thing in railway control, and another thing in wire-transmitted telecommunications, leading to two definitions, one for *signal (railways)* and a different one for *signal (telecommunications).* The specialisations in electrotechnical terminology usually follow the designations of the IEC Technical Committees. TC 9 is Electrical equipment and systems for railways); there are many TCs which could (and do) use a telecommunications notion (in fact, signal has many definitions; see 5.1). The second way is by reconciling the two different definitions into one. Considerations towards the second path are illustrated here.

Much of both definitions is the same, but some of it is definitively different. The IACS in question is uniquely determined: it is whichever system the IEC 62443 standard is applied to in the instance of its application. The organisation involved (according to the first definition) might also be unique, but it could be that many organisations are involved in the joint ownership or custodianship of an asset. Is an IACS, as a non-sentient physical object, an object of which it might make any sense at all to speak of as having values? Or is the valuer an implicit organisation which is considering *Ob* and *theIACS* together, to determine whether there is a perceived or actual “value” (causal influence?) of the one on the other? Say, *PV(Ob,theIACS,Org) OR AV(Ob,theIACS,Org).* The SemAn analyst cannot decide such matters; the domain specialists writing the standard must do so.

5.3.4 Output of SemAn Analyser on *asset*

There are two non-identical definitions of asset in the IEC 62443 series of standards. Here is the output of the SemAn Analyser on both:

```
10.
asset:

physical or logical object owned by or under the custodial duties of an
organization, having either a perceived or actual value to the organization

\\
     physical or logical object
          > [AND] owned by or under the custodial duties
                       > of an organization
          > [AND] having either a perceived or actual value
                       > to the organization
[Source: IEC 62443-2-1:2010]
[Source: IEC TS 62443-1-1:2009]

11.
asset:

physical or logical object having either a perceived or actual value to the
IACS

\\
     physical or logical object
          > having either a perceived or actual value
                  > to the IACS
[Source: IEC 62443-3-3:2013].
```

This annotated parsing/pretty-printing shows immediately and clearly the similarities and differences immediately which we have recognised in the more laborious manual SemAn. Namely, first, an *asset* is a *physical or logical object*; and, second, this object *[has] a perceived or actual value.* To whom the value accrues is clearly different in the two cases. Further, one definition mentions custodial duties associated with the asset; the other does not. This comparison gives clear indications of the differences seen during the manual SemAn, and leads to the same scope of discussion for domain experts attempting to harmonise the two definitions as did the manual analysis.

5.3.5 Conclusions

Experience has shown that use of the SemAn Analyser indeed eases the task of performing a SemAn, as it clearly did in the case of *asset* in 5.4.4 and 5.4.5. In the case of *harm*, 5.4.3 showed that there were many meaning postulates that played a role in eliminating/reducing some of the terms occurring in the definition, which the SemAn Analyser treats as primitive. So here the manual SemAn achieved results which the SemAn Analyser could not obtain.

5.4 HintAn and HintAn Analyser

5.4.1 General

The form of definitions of terminology in IEC standards is constrained by ISO/IEC Directives (especially Part 2 Section 16) as well as by the terminology standard IEC 10241-1:2011 (confirmed 2022). There are many definitions in IEC standards which do not conform with these constraints. HintAn selects some of these constraints, and the HintAn Analyser checks whether the constraints are satisfied on an input collection of definitions, and outputs locations which are violations. The constraints are listed in 5.4.2 and discussed individually with examples in 5.5.

The collection of 19 constraints forming HintAn and partially implemented in the HintAn Analyser can be classified into four categories.

5.4.2 Usability

This category speaks to the ease with which a definition may be understood and thus used in standards. The length of a definition and its structural complexity do not make a definition unusable, but a definition which is hard cognitively to grasp may be more easily misused than one which is easier to understand. The two constraints are:

- Definition is Comparatively Long,
- Definition is Structurally Complex.

The comparative length of a definition is assessed through its percentile in the collection of all definitions in the terminology database. The structural complexity of a definition is assessed through the output of the SemAn Analyser. These are hints to the formulators of the definition which may or may not be followed, depending on the justification for the choice of definiens.

5.4.3 Ill-Formedness

Some definitions are syntactically ill-formed. The syntactic form of definitions is explicated in ISO/IEC Directives Part 1: 2021 Annex SK and in IEC 10241-1:2011. The four constraints associated with ill-formedness are:

- Too many definitions.
- Contains sentences.
- No Definition.
- with subcategory – Definition Contains only a Formula.
- Definition Contains Field of Use without Angle Brackets.
- Reference Not in Appropriate Form.

A definition should contain one definiens. If it contains more than one, there are “*too many definitions*”. If there are many concepts associated with a term, then the term is homonymous, which is deprecated in IEC Terminology. One solution is to specify a *field of use* for each non-equivalent definiens. Then there will be as many definitions of the term as there are non-equivalent definiens, each qualified with a (non-overlapping) field of use. An example was given for *signal* in 5.2.

The substitutability requirement (see Secion 5.4.4) entails that a definiens cannot contain a sentence (or of course more than one). This is classified here as an ill-formedness condition, but it could equally well be classified as a failure of substitutability.

The definiens may be absent (without a reference being specified, as below), or it may be syntactically ill-formed (which can be determined from the output of the SemAn Analyser). In this case, it is designated "*no definition*" by HintAn.

The Directives Part 2 specify that a field of use is to be denoted within "angle brackets", e.g. <*railway systems*>.

Similarly, a definition may be given in a standard that contains no explicit definiens, but instead a reference to a subclause of another standards document in which the definition of the term is given. Such a reference must be given in "square brackets" and must contain the document/subclause reference in standard IEC form, e.g. *[IEC 61508-4:2010 3.1.4]* (which is the definition of *hazardous event*).

5.4.4 Violates Substitutability

It is intended that a definiens be able syntactically to substitute for its definiendum (the term omitted, and the definiens put in its place) everywhere in a standard while retaining the original meaning of the sentence or phrase in which the substitution has taken place. This requirement has some formal consequences:

- if the definiens commences with a determiner ("a", "the") then, after substitution, there will likely be two consecutive determiners in the sentence – an original determiner preceding the term, and the second introduced in the definiens. The sentence or phrase will be syntactically ill-formed.
  Example:

  Phrase: "The cow jumped over the moon."

  Definition: Cow = a female bovine.

  Substitution: "The a female bovine jumped over the moon."
- if the definiens terminates with a full stop, then after substitution this will terminate the sentence at the end of the definiens, and the rest of the original sentence or phrase will be left hanging.
  Example:

  Phrase: "The cow jumped over the moon."

  Definition: cow = female bovine.

  Substitution: "The female bovine. jumped over the moon."
- if the definiens commences with an upper-case letter, then the resultant substitution is awkward, since a upper-case letter in English that does not commence a sentence is usually only used for a specific named object
  Example:

  Phrase: "The cow jumped over the moon."

  Definition: Cow = Female bovine.

  Substitution: "The Female bovine jumped over the moon."
- if the definiens contains the definiendum ("Cyclical Definition"), then it is not possible to replace the term with the definiens – it will still occur after the replacement. The simplest, maybe the most common, case of this is:

  Phrase: "The cow jumped over the moon."

  Definition: Cow = cow is a female bovine.

  Substitution: "The cow is a female bovine jumped over the moon."

The categories associated with a violation of substitutability are:

- Term cannot be replaced by the Definiens.
- Definition Starts with a Determiner.

- Definition Terminates with Full Stop.
- Definition Starts with Upper-Case letter.
- Cyclical Definition.

5.4.5 Inappropriate Definiens

- Negative Form.
- Definition Contains a Requirement.
- Definition Contains Items Which Should be in Notes.
- Constructional Definition.
- Definition Contains Superfluous Phrase.
- Definition Introduces Synonym or Hyperonym.
- Definition Addresses Use but not Concept.

These categories are explained in more detail, with examples, in 5.5.

5.5 HintAn Categories by Example

5.5.1 Definition is Comparatively Long

The examples below lie in the top 1% of lengths in the existent DKE database. This means that there exist definitions of this length already. It is a matter of judgement whether this is appropriate in any given case.

Table 1 – Examples for comparatively long definitions

publication_id	term	label	def
IEC TR 80001-2-3:2012	medical device	3.41	any instrument, apparatus, implement, machine, appliance, implant, in vitro reagent or calibrator, software, material or other similar or related article: a)) intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the specific purpose(s) of: – diagnosis, prevention, monitoring, treatment or alleviation of disease, – diagnosis, monitoring, treatment, alleviation of or compensation for an injury, – investigation, replacement, modification, or support of the anatomy or of a physiological process, – supporting or sustaining life, – control of conception, – disinfection of medical devices, – providing information for medical or diagnostic purposes by means of in vitro examination of specimens derived from the human body; and b)) which does not achieve its primary intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its intended function by such means.
IEC TS 62556:2014	transition distance	3.100	for a given longitudinal plane, for design: the transducer aperture area of the ultrasonic transducer divided by π times the effective wavelength λ, i.e. ATA/(πλ); for measurements: the transducer aperture area is replaced by the source aperture area so that the transition distance is given by ASAeff/(πλ). For design, for an unapodized ultrasonic transducer with circular symmetry about the beam axis, the source aperture area is πae2, where ae is the effective radius; therefore, the transition distance is zT = ae2/λ. For design, for an unapodized rectangular ultrasonic transducer which has a transducer aperture width, LTA1, in a specified longitudinal plane, the

publication_id	term	label	def
			effective in-plane area is (LTA1)2. Therefore, for this plane, the transition distance is zT1 = (LTA1) 2 /(πλ). The transition distance for the orthogonal longitudinal plane including the second transducer aperture width is zT2 = (LTA2) 2 /(πλ). Similarly, for measurements in each specified longitudinal plane, the source aperture width in that plane is used, or zT1 = (LSA1)2/(πλ); in the other orthogonal plane, zT2 = (LSA2)2 /(πλ).
IEC TS 62862-1-1:2018	levelized electricity cost	3.10.1	cost of the net electricity produced by the STE plant, usually given in and calculated using the equation: |FORMULA|, |FORMULA|, Cinvest is the total plant investment [Unit: $ or €], C(O&M) is the annual operation and maintenance costs [Unit: $ or €], Cfuel is the annual fuel costs [Unit: $ or €], Enet is the annual net plant electricity production, rd is the real debt interest rate [Unit: dimensionless, usually expressed as a percentage, %], n is the depreciation period [Unit: years], rinsurance is the annual insurance rate, given as % of Cinvest.
IEC/IEEE 60780-323:2016	equipment important to safety	3.12	equipment that is part of a safety group and/or whose malfunction or failure could lead to undue radiation exposure of the site personnel or members of the public. Equipment including: • those structures, systems and components that prevent anticipated operational occurrences from leading to accident conditions; • those features that are provided to mitigate the consequences of malfunction or failure of structures, systems, and components. A)) For usage consistent with IEC 61226, equipment important to safety are as follows: – all I&C equipment performing Category A to Category C functions (in accordance with the IEC 61226 categorisation scheme), - all electrical equipment needed to ensure emergency energy supply to this equipment in case of a loss of normal power supply, – all electrical equipment needed to ensure ultimate energy supply in case of total loss of on-site power (if selected as design extension condition to be mitigated). B)) For usage consistent with other IEEE documents and a Class 1E categorization; for equipment important to safety, qualification is essential to the following: – electric equipment and systems that are essential to emergency reactor shutdown, containment isolation, reactor core cooling, and containment and reactor heat removal, or – electric equipment that are otherwise essential in preventing significant release of radioactive material to the environment.

5.5.2 Definition is Structurally Complex

The following definition, output from the SemAn Analyser, contains logical constants (here, *AND* and *OR*) “nested” to depth 4. Amongst scientists who deal with complex logical formulas, the depth of nesting in formulas is a standard indication of cognitive complexity.

```
43.
configuration baseline:
```

```
information that allows the software release to be recreated in an auditable and
systematic way, including: all source code, data, run time files, documentation,
configuration files, and installation scripts that comprise a software release;
information about compilers, operating systems, and development tools used to create
the software release

\\
information
   > that allows the software release
      > [AND] to be recreated in an auditable and systematic way
      > [AND] , including: all | [AND] source code
                              | [AND] data
                              | [AND] information
                              |            [AND] about compilers
                              |            [AND] , operating systems
                              |            [AND] , and development tools
                              |             > used to create the software release
                              |             > [AND] , run time files
                              |             > [AND] , documentation
                              |             > [AND] , configuration files
                              |             > [AND] , and installation scripts
                              |             > that comprise a software release
[Source: IEC 61508-4:2010].
```

In this particular case, the nesting is caused by the need to enumerate specific categories of object in various places the predicate. Such listing is not necessarily cognitively complex. It is for the proposer of the definition, as well as the standards user, to judge this aspect. Structural complexity is a matter of judgement.

5.5.3 Too Many Definitions

According to IEC Directives Part 2, there should be precisely one definiens per term/definiendum. If a term has many non-equivalent meanings (i.e., it is a homonym), these meanings can be distinguished by assigning a field of use to each. The fields of use are typically assigned in conformance with subject matter of the IEC Technical Committee responsible for the specific standard in which the term is defined.

Examples:

Table 2 – Examples for terms with too many definitions

publication_id	term	label	def
IEC TS 62351-2:2008	worm	2.2223	1. A computer program that can run independently, can propagate a complete working version of itself onto other hosts on a network, and may consume computer resources destructively. 2. A computer virus capable of disrupting a computer program. 3. A self-contained program that can propagate itself through systems or networks. 4. [An] independent program that replicates from machine to machine across network connections, often clogging networks and computer systems as it spreads.
IEC 60092-201:2019	steering gear power unit	3.1.9	a)) in the case of electric steering gear, an electric motor, and its associated electrical equipment; b)) in the case of electrohydraulic steering gear, an electric motor and its associated electrical equipment and connected pump.
IEC 61158-5-11:2007	group	3.6.62	a)) a general term for a collection of objects. Specific uses: b)) when describing an address, an address that identifies more than one entity.
IEC 61512-1:1997	batch	3.5	(1) The material that is being produced or that has been produced by a single execution of a batch process. (2) An entity that represents the production of a material at any point in the process.

5.5.4 Contains Sentences

A definiens should not contain a sentence or multiple sentences.

Table 3 – Examples for definitions with one or multiple sentences

publication_id	term	label	def
IEC TR 62051-1:2004	registered COSEM names (relevant to IEC 62056-53 and IEC 62056-62)	3.1.104	items used for the COSEM meter model and DLMS/COSEM protocols, which must have globally unique and unambiguous names. These objects are registered by the DLMS user association and include the following: • COSEM interface class_id-s and versions; • standard COSEM object identifiers (OBIS names); • COSEM_Application_Context_Names; • COSEM_Authentication_Mechanism_Names.
IEC TR 62051-1:2004	management logical device (relevant to IEC\xa062056 -62)	3.1.71	mandatory element of any physical device, with a reserved address. It must support an application association to a public client with the lowest security level. Its role is to support the revelation of the internal structure of the physical device and the notification of events in the server.

These examples are also used as examples in 5.5.1, for they also violate the condition that a definition shall not contain a requirement.

5.5.5 No Definition

There are various ways in which a definiens may be considered to be missing. One is legitimate, namely a reference (in standard form) to a definition in another standard. The following are illegitimate:

- the definition does not contain a definition text in the narrower sense or a reference to another entry.
- Occasionally: The definition text is completely missing. Only a term is given, potentially directly followed by notes or examples.
- Occasionally: The definition text only contains text in brackets, which is not a reference to another entry.

Table 4 – Examples for missing definitions

publication_id	term	label	def
IEC 60601-1-10, ed. 1.0 (2007-11)	E	3.6	
IEC 62728, ed. 1.0 (2011-09)	quantum efficiency	4.2.8	
IEC 61009-1, ed. 3.0 (2010-02)	non-operating overcurrents in the main circuit	3.4.19	NOTE In the case of overcurrent in the main circuit, in the absence of residual current, operation of the detecting device may occur as a consequence of asymmetry existing in the detecting device itself.
IEC 61427-2, ed. 1.0 (2015-08)	accuracy	3.1	<of a measuring instrument>
IEC/TS 61836, ed. 3.0 (2016-12)	photovoltaic	3.1.43	(Abbreviation: PV)

A subcategory of this, for which HintAn provides a separate output hint, is that the definiens contains only a formula.

Table 5 – Example for missing definition containing only a formula

publication_id	term	label	def
IEC 60876-1, ed. 5.0 (2014-08)	logarithmic transfer matrix	3.1.4	aij = –10 log10 ti

5.5.6 Definition Contains Field of Use Without Angle Brackets

In these cases, the field of use is specified in running text inside the definiens. This is not conformant with the IEC Directives. A field of use is to be specified next to the term, in “angle brackets”. Examples and “corrected” examples are given (the examples are deficient in other ways, and some of these other ways have been corrected in the “corrected” example, but not necessarily satisfactorily. In these cases, the definitions can be referred back to the originating TC for improvement).

Table 6 – Examples for definitions containing field of use without angel brackets

publication_id	term	label	def
IEC 60068-2-80:2005	standard deviation	3.32	in vibration theory, the mean value of vibration is equal to zero. Therefore for a random time history, the standard deviation is equal to the r.m.s. value.
IEC 60076-6:2007	rated reactance	9.3.15	for damping and discharge reactors, the specified reactance of the reactor at power frequency. The rated reactance is expressed in Ohms per phase.
IEC 61076-4-108:2002	standard equipment	2.1.2.3	full grid:contact arrangement in which all grid positions are equipped. chess pattern grid:contact arrangement in which the grid positions are equipped alternately like the squares on a chess board.
IEC 62055-31:2005	specified matching socket	3.1.9	in relation to a payment meter arranged as a plug-in unit, a specified matching socket comprises a base with jaws to accept and connect to the plug-in unit, terminals for connection of the supply network and the consumer load circuit, and appropriate secure fixing and sealing arrangements. The payment meter is capable of meeting the relevant type-testing requirements when it is properly installed in any specified matching socket.
IEC 62056-2:2017	enabled	3.5.7	when used in the context of “Credit” or “Charge” types; means that the “Credit” or “Charge” type appears in the credit_reference_list or charge_reference_list respectively of the “Account” object.

Table 7 – Corrected examples from Table 6

publication_id	term	label	def
IEC 60068-2-80:2005	standard deviation <vibration theory>	3.32	r.m.s. value
IEC 60076-6:2007	rated reactance <damping and discharge reactors>	9.3.15	specified reactance of the reactor at power frequency, expressed in Ohms per phase
IEC 61076-4-108:2002	Standard equipment <full grid>		contact arrangement in which all grid positions are equipped

publication_id	term	label	def
IEC 61076-4-108:2002	standard equipment <chess pattern grid>	2.1.2.3	contact arrangement in which the grid positions are equipped alternately like the squares on a chess board
IEC 62055-31:2005	specified matching socket <payment meter as plug-in>	3.1.9	base with jaws to accept and connect to the plug-in unit, terminals for connection of the supply network and the consumer load circuit, and appropriate secure fixing and sealing arrangements.
IEC 62056-6-2:2017	Enabled <"Credit" or "Charge" type>	3.5.7	"Credit" or "Charge" type appears in the credit_reference_list or charge_reference_list respectively of the "Account" object.

5.5.7 Reference Not in Appropriate Form

Examples of correct references are:

When the source is in the IEV:

SOURCE: IEV 192-03-01:2015-02, modified – Note 1 to entry has been omitted

When the source is in a standards document:

[SOURCE: IEC 61508-4:2010 3.1.4]

Any reference which is formatted other than these is not in the appropriate form.

5.5.8 Term Can Not Be Replaced by the Definition

ISO/IEC Directive Part 1 and Supplement:2021 Annex SK explains this condition as follows:

> "The definition **shall have the same grammatical form as the term**. Thus, to define a verb, a verb shall be used; to define a noun in the singular, the singular shall be used. In the case of adjectives, it is often essential to indicate in the definition to which objects the concept applies. The definition then begins with "qualifies" or "pertaining to" (see ISO/IEC Directive Part 1 and Supplement:2021, Annex SK.3.1.3.6.2). In the case where a term is an adjective and a noun (see ISO/IEC Directive Part 1 and Supplement:2021, Annex SK.3.1.3.6.2), the definition shall be written so as to apply to both parts of speech".

Also,

> "The form of a definition shall be such that it **can replace the term in the context where the term appears**" (see ISO/IEC Directive Part 1 and Supplement:2021, Annex SK.3.1.4.2).

Table 8 – Examples for terms which can't be replaced by their definition

publication_id	term	label	def
IEC 60044-8:2002	phase error	3.1.29	The phase error is the phase displacement φ minus the displacement caused by the rated phase offset and the rated delay time. The phase error is related to the rated frequency. |FORMULA| The phase error is usually expressed in minutes or centiradians.
IEC 60122-1:2002 + AMD 1:2017 CSV	drive level dependency	2.2.37	drive level dependency (DLD) is the effect of changes in drive level conditions upon the resonance resistance of the crystal unit. This parameter can be specified by defining the ratio of resistance between two specified drive levels. This ratio is represented by the expression: |FORMULA|

publication_id	term	label	def
			|LIST|
IEC 60534-9:2007	static	3.18	means without motion or change [4]; readings are recorded after the device has come to rest. Static performance can be measured either without process loading (bench-top tests), with simulated or active loading, or under process operating conditions
IEC 60761-1:2002	Units	3.26	This standard uses the SI system of units. The following units of practical importance are also used where appropriate: – for time: year (y), day (d), hour (h), minute (min), – for energy: electron-volt (eV), – for volumic activity: Becquerel per cubic metre (Bq/m3).
IEC 62288, ed. 2.0 (2014-07)	alarm	3.4	(MSC.302/A) a high-priority alert. Condition requiring immediate attention and action by the bridge team, to maintain the safe navigation of the ship.

Table 9 – Corrected examples from Table 8

publication_id	term	label	def
IEC 60044-8:2002	phase error	3.1.29	phase displacement φ minus the displacement caused by the rated phase offset and the rated delay time, usually expressed in minutes or centiradians
IEC 60122-1:2002 + AMD 1:2017 CSV	drive level dependency	2.2.37	effect of changes in drive level conditions upon the resonance resistance of the crystal unit Note to Entry: This parameter can be specified by defining the ratio of resistance between two specified drive levels. This ratio is represented by the expression: |FORMULA| |LIST|.
IEC 60534-9:2007	static	3.18	without motion or change Note 1 to Entry. Readings are recorded after the device has come to rest. Note 2 to Entry. Static performance can be measured either without process loading (bench-top tests), with simulated or active loading, or under process operating conditions.
IEC 60761-1:2002	Units	3.26	SI units Note to Entry. The following units of practical importance are also used where appropriate: – for time: year (y), day (d), hour (h), minute (min), – for energy: electron-volt (eV), – for volumic activity: Becquerel per cubic metre (Bq/m3).
IEC 62288, ed. 2.0 (2014-07)	alarm	3.4	high-priority alert requiring immediate attention and action by the bridge team to maintain safe navigation

5.5.9 Definition Starts with a Determiner

Explained with brief example in 5.4.4. Examples above contain such occurrences, for example in 5.5.3 and 5.5.7. This feature has been corrected in the “corrected” examples in 5.5.7.

Note on Table 10: The example in Table 10 is also an example of containing a sentence as well as ending a definition with a full stop

Table 10 – Example of definitions starting with a determiner

publication_id	term	label	def
CISPR 16-2-4:2003	Unsymme-trical mode (V-terminal voltage)	3.12	the voltage between a conductor or terminal of a device, equipment or system and a specified ground reference. For the case of a two-port network, the two unsymmetrical voltages are given by: a)) the vector sum of the asymmetrical voltage and half of the symmetrical voltage; and b)) the vector difference between the asymmetrical voltage and half of the symmetrical voltage.

Table 10 – Corrected Example from Table 10

publication_id	term	label	def
CISPR 16-2-4:2003	Unsymme-trical mode (V-terminal voltage)	3.12	voltage between a conductor or terminal of a device, equipment or system and a specified ground reference

5.5.10 Definition Terminates with a Full Stop

Explained with a brief example in 5.4.4 and above in 5.5.8.

5.5.11 Definition Starts with Upper-Case Letter

Explained with a brief example in 5.4.4.

5.5.12 Cyclical Definition

When the term/definiendum occurs in the definiens, it is not possible to eliminate the term by a phrase which does not include the term, since the replacement, which is the definiens, still includes the term.

Table 11 – Examples of cyclical definitions

publication_id	term	label	def
IEC 61158-6-2, ed. 3.0 (2014-08)	resource	3.3.59	resource is a processing or information capability of a subsystem
ISO/IEC 14165-222, ed. 1.0 (2005-05)	device	3.1.10	in this document, the term device is used to refer to an I/O device such as a printer, magnetic-tape unit and direct-access-storage device. The I/O device operation is regulated by a control unit that provides the logical and buffering capabilities necessary to operate the I/O device
IEC 62823, ed. 1.0 (2015-08)	line current	3.13	power frequency line current
ISO/IEC 14543-4-1, ed. 1.0 (2008-05)	node	3.1.21	communication node conforming to ISO/IEC 14543-4;

Note on Table 13: Here, in one example it seems as if the definiendum *line current* is intended to be an abbreviation of a longer term *power frequency line current*. Abbreviations are handled in standards documents differently from this, so the “correction” to be made is other than altering the definiens.

Table 12 – Corrected examples from Table 12

publication_id	term	label	def
IEC 61158-6-2, ed. 3.0 (2014-08)	resource <subsystem>	3.3.59	processing or information capability
ISO/IEC 14165-222, ed. 1.0 (2005-05)	device	3.1.10	I/O device such as a printer, magnetic-tape unit and direct-access-storage device Note to Entry. The I/O device operation is regulated by a control unit that provides the logical and buffering capabilities necessary to operate the I/O device.
ISO/IEC 14543-4-1, ed. 1.0 (2008-05)	node	3.1.21	communication device conforming to ISO/IEC 14543-4

5.5.13 Negative Form

The definiens attempts to explain the term by stating what it is *not* or what it *does not contain*. Such a definiens should be replaced if possible by one which says what it is, or what it contains. This is a matter for expert domain knowledge.

Table 13 – Examples of negative form definitions

publication_id	term	label	def
CISPR 12, ed. 6.0, amd. 1 (2009-01)	device	3.3	machine driven by an internal combustion engine which is not primarily intended to carry persons or goods
IEC 62676-1-1, ed. 1.0 (2013-10)	normal operation	3.1.95	state of the VSS when not in power-up or power down procedures and no fault is present
IEC 62718, ed. 1.0 (2013-04)	rated voltage	3.1.4	voltage declared by the manufacturer to which all the electronic ballast characteristics are related and which is not less than 85 % of the maximum value of the rated voltage range
IEC 62128-1, ed. 2.0 (2013-09)	leakage current	3.6.2	current which, in the absence of a fault, flows to earth or to extraneous conductive parts in a circuit
ISO/IEC 14776-151, ed. 1.0 (2010-07)	direct current (D.C.)	3.1.46	the non-A.C. component of a signal

Note on Table 15: To correct these definitions expert judgement is required. These modifications are proposals as examples in this document and have not been modified by a domain expert in the responsible IEC TC. Also note that the definition of “normal operation” stays the same. It is apparently one of four mutually-exclusive system states, of which the other three have special definitions.

Table 14 – Corrected examples from Table 14

publication_id	term	label	def
CISPR 12, ed. 6.0, amd. 1 (2009-01)	device	3.3	machine driven by an internal combustion engine which is other than a personnel or goods vehicle
IEC 62676-1-1, ed. 1.0 (2013-10)	normal operation	3.1.95	state of the VSS when not in power-up or power down procedures and no fault is present

publication_id	term	label	def
IEC 62718, ed. 1.0 (2013-04)	rated voltage	3.1.4	voltage declared by the manufacturer to which all the electronic ballast characteristics are related, and which is greater or equal to 85 % of the maximum value of the rated voltage range
IEC 62128-1, ed. 2.0 (2013-09)	leakage current	3.6.2	current which flows to earth or to extraneous conductive parts in a circuit in normal operation
ISO/IEC 14776-151, ed. 1.0 (2010-07)	direct current (D.C.) <signal>	3.1.46	complement to the A.C. component

5.5.14 Definition Contains a Requirement

“A definition shall not take the form of, or contain, a requirement.” (ISO/IEC DIR 1:2021 + IEC SUP:2021, Annex SK; p.114)

There are “trigger words” which indicate a formal requirement: “must”, “shall”,

Table 15 – Examples of definitions containing a requirement

publication_id	term	label	def
IEC 60870-6-503, ed. 3.0 (2014-07)	operation	3.12	activity which shall be performed by the TASE.2 server at the request of the TASE.2 client
IEC 61019-1, ed. 1.0 (2004-11)	operating temperature range	4.2.4	range of temperatures as measured on the enclosure over which the resonator must function within the specified tolerances
IEC TR 62051-1:2004	registered COSEM names (relevant to IEC\xa062056 -53 and IEC\xa062056 -62)	3.1.104	items used for the COSEM meter model and DLMS/COSEM protocols, which must have globally unique and unambiguous names. These objects are registered by the DLMS user association and include the following: • COSEM interface class_id-s and versions. • standard COSEM object identifiers (OBIS names), • COSEM_Application_Context_Names, • COSEM_Authentication_Mechanism_Names.
IEC TR 62051-1:2004	manage-ment logical device (relevant to IEC\xa062056 -62)	3.1.71	mandatory element of any physical device, with a reserved address. It must support an application association to a public client with the lowest security level. Its role is to support the revelation of the internal structure of the physical device and the notification of events in the server
IEC PAS 62282-6-150:2011	fuel byproducts	3.40	Class 8 (corrosive), or non-hazardous compounds produced during the generation of hydrogen and/or electricity from solid water reactive fuel; fuel byproducts shall not have any subsidiary risks.

Note on Table 17: Requirements, such as that *fuel byproducts shall not have any subsidiary risks* belong, not in Clause 3 Terms and Definitions, but in subsequent normative clauses of the standard. Correcting these examples is a matter for domain experts; however, proposals are offered as follows:

Table 16 – Corrected examples from Table 16

publication_id	term	label	def
IEC 60870-6-503, ed. 3.0 (2014-07)	operation	3.12	activity which it to be performed by the TASE.2 server at the request of the TASE.2 client

publication_id	term	label	def
IEC 61019-1, ed. 1.0 (2004-11)	operating temperature range	4.2.4	range of temperatures as measured on the enclosure over which the resonator is intended to function within specified tolerances
IEC TR 62051-1:2004	registered COSEM names (relevant to IEC\xa062056 -53 and IEC\xa062056 -62)	3.1.104	items used for the COSEM meter model and DLMS/COSEM protocols
IEC TR 62051-1:2004	management logical device (relevant to IEC\xa062056 -62)	3.1.71	mandatory element of any physical device, with a reserved address
IEC PAS 62282-6-150:2011	fuel byproducts	3.40	Class 8 (corrosive), or non-hazardous compounds produced during the generation of hydrogen and/or electricity from solid water reactive fuel.

5.5.15 Definition Contains Items That Should Be in Notes/Examples

Self-explanatory category. There are “trigger words” such as *e.g.*, *such as*, *typically* which indicate the presence of examples. There is “trigger punctuation”, in that phrases which should be in Notes or Examples are often expressed in separate sentences or apposite phrases. This category, then, overlaps with 5.5.3 Too Many Definitions.

Table 17 – Examples for definitions that contain items which should be in notes/examples

publication_id	term	label	def
CISPR 16-2-4:2003	Unsymmetrical mode (V-terminal voltage)	3.12	the voltage between a conductor or terminal of a device, equipment or system and a specified ground reference. For the case of a two-port network, the two unsymmetrical voltages are given by: a)) the vector sum of the asymmetrical voltage and half of the symmetrical voltage; and b)) the vector difference between the asymmetrical voltage and half of the symmetrical voltage.
CISPR 20:2006 + AMD 1:2013 CSV	sound receivers	3.1.1	appliances intended for the reception of sound broadcast and similar services for terrestrial, cable and satellite transmissions; these sound receivers can be digital receivers with digital incoming signals or receivers with digital processing of digital or analogue incoming signals
CISPR 20:2006 + AMD 1:2013 CSV	disturbance signal	3.1.5	an unwanted signal which may degrade radio reception or cause malfunction in equipment; specific unwanted signals are simulating disturbance signals, generated under laboratory conditions
IEC 60034-2-3:2020	switching frequency	3.5	number of switching events of one semiconductor within one second. It determines, together with the selected pulse pattern and the converter topology, the lowest frequency of non-controllable high frequencies or inter-harmonics at the IPC (in-plant point of coupling) or the motor
IEC 60076-6:2007	rated power	7.3.3	reactive power of the shunt reactor specified for operation at rated voltage and rated frequency. See also note in 7.4.3. In the case of reactors with adjustable reactance, the rated power refers to the setting of the reactor with highest reactive power, unless otherwise specified.

Table 18 – Corrected examples from Table 18

publication_id	term	label	def
CISPR 16-2-4:2003	Unsymmetrical mode (V-terminal voltage)	3.12	voltage between a conductor or terminal of a device, equipment or system and a specified ground reference Note to Entry. For the case of a two-port network, the two unsymmetrical voltages are given by: a)) the vector sum of the asymmetrical voltage and half of the symmetrical voltage; and b)) the vector difference between the asymmetrical voltage and half of the symmetrical voltage.
CISPR 20:2006 + AMD 1:2013 CSV	sound receivers	3.1.1	appliances intended for the reception of sound broadcast and similar services for terrestrial, cable and satellite transmissions Note to Entry. These can be digital receivers with digital incoming signals or receivers with digital processing of digital or analogue incoming signals
CISPR 20:2006 + AMD 1:2013 CSV	disturbance signal	3.1.5	unwanted signal which may degrade radio reception or cause malfunction in equipment Example. specific unwanted signals are simulating disturbance signals, generated under laboratory conditions.
IEC 60034-2-3:2020	switching frequency	3.5	number of switching events of one semiconductor within one second Note to Entry. It determines, together with the selected pulse pattern and the converter topology, the lowest frequency of non-controllable high frequencies or inter-harmonics at the IPC (in-plant point of coupling) or the motor.
IEC 60076-6:2007	rated power	7.3.3	reactive power of the shunt reactor specified for operation at rated voltage and rated frequency Note 1 to Entry. See also note in 7.4.3. Note 2 to Entry. In the case of reactors with adjustable reactance, the rated power refers to the setting of the reactor with highest reactive power, unless otherwise specified.

5.5.16 Constructional Definition

A constructional definition is a definiens which defines through enumerating parts of the definiendum rather than through a description of function. It is sometimes appropriate, but the IEC Directives indicate using a description of function in preference.

Table 19 – Examples of constructional definitions

publication_id	term	label	def
IEC 62110:2009	power system	3.7	system consisting of overhead lines and underground cables, substations and other power distribution and transmission equipment. Railway systems are covered by a specific standard and therefore are excluded from the present standard.
IEC 62134-1:2009	cabinet	3.10	container that may enclose connection devices, terminations, apparatus, cabling, and equipment. It may be either wall-mounted or self-supporting
IEC 62382:2012	basic software	3.1.9	software which, at a minimum, contains the graphic faceplates, base-level alarms and switch points, basic interlocking, and analogue control. In the case of safety loops, any safety switch point should be included if it is not in the basic database

publication_id	term	label	def
IEC 62496-2-2:2011	optical position adjusting system	3.1	consists of a light source, fibre position adjustment stage, OCB holder, input/output fibre and a power meter. The optimum fibre launch position, at which the optical output power is maximised, is determined through alignment of the input/output fibres to the OCB and monitoring the output power from the OCB
IEC 62520:2011	primary	3.21	the primary comprises three parts: a three-phase winding, a slotted laminated ferromagnetic core, and a mechanical support structure

Table 20 – Corrected examples from Table 20

publication_id	term	label	def
IEC 62110:2009	power system	3.7	system of overhead lines and underground cables, substations and other power distribution and transmission equipment Note to Entry. Railway systems are covered by a specific standard and therefore are excluded from the present standard.
IEC 62134-1:2009	cabinet	3.10	container Note 1 to Entry. It may enclose connection devices, terminations, apparatus, cabling, and equipment. Note 2 to Entry. It may be either wall-mounted or self-supporting.
IEC 62382:2012	basic software	3.1.9	software containing graphic faceplates, base-level alarms and switch points, basic interlocking, and analogue control Note to Entry. In the case of safety loops, any safety switch point should be included if it is not in the basic database.
IEC 62520:2011	primary	3.21	item with three-phase winding, slotted laminated ferromagnetic core and mechanical support structure

5.5.17 Definition Contains Superfluous Phrase

Self-explanatory

Table 21 – Examples of definitions which contains superfluous phrases

publication_id	term	label	def
IEC 60068-2-20, ed. 5.0 (2008-07)	contact angle	3.2	in general, the angle enclosed between two planes, tangent to a liquid surface and a solid/liquid interface at their intersection (see Figure 1). In particular the contact angle of liquid solder in contact with a solid metal surface
IEC 62564-1, ed. 2.0 (2011-08)	component	3.1.6	either a microcircuit, integrated circuit, semiconductor, or discrete semiconductor for the purpose of this specification
IEC 61000-3-6, ed. 2.0 (2008-02)	fundamental frequency	3.26.1	frequency in the spectrum obtained from a Fourier transform of a time function, to which all the frequencies of the spectrum are referred. For the purpose of this technical report, the fundamental frequency is the same as the power supply frequency
IEC 62368-2, ed. 2.0 (2015-02)	double insulation	3.3.5.2	Purpose: To support the concept of safeguards as used in this standard.
IEC 61557-8, ed. 3.0 (2014-12)	overload current	3.1.24	overload current occurring in an electric circuit according to this standard is overload current which is caused by connected loads

Table 22 – Corrected examples from Table 22

publication_id	term	label	def
IEC 60068-2-20, ed. 5.0 (2008-07)	contact angle	3.2	angle enclosed between two planes, tangent to a liquid surface and a solid/liquid interface at their intersection Note to Entry. In particular, the contact angle of liquid solder in contact with a solid metal surface.
IEC 62564-1, ed. 2.0 (2011-08)	component	3.1.6	microcircuit, integrated circuit, semiconductor, or discrete semiconductor
IEC 61000-3-6, ed. 2.0 (2008-02)	fundamental frequency	3.26.1	frequency in the spectrum obtained from a Fourier transform of a time function, to which all the frequencies of the spectrum are referred Note to Entry. In this document, the fundamental frequency is the same as the power supply frequency.
IEC 62368-2, ed. 2.0 (2015-02)	double insulation	3.3.5.2	[no correction offered]
IEC 61557-8, ed. 3.0 (2014-12)	overload current	3.1.24	[no correction offered]

5.5.18 Definition Introduces Synonym/Hyperonym

A putative definition which introduces another term or terms with the same or similar meaning is not an appropriate definition according to IEC Directives. Such examples can be easily modified: according to the standard formatting of Clause 3, synonyms are to be listed as definienda; then the definiens can be given as a reference to a source document elsewhere, or to the IEV.

Examples:

Table 23 – Examples of definitions introducing synonyms/hyperonyms

publication_id	term	label	def
IEC 62390, ed. 1.0 (2005-01)	resource	3.1.23	– logical device
IEC 60987, ed. 2.0 (2007-08)	FTA	3.6	fault tree analysis
IEC 61158-6-10, ed. 3.0 (2014-08)	message	3.2.68	synonym for frame
IEC 62065, ed. 2.0 (2014-02)	alarm	3.1.2	high-priority alert
IEC 61158-3-13, ed. 2.0 (2014-08)	application process	3.3.1	application layer task

Table 24 – Corrected examples from Table 24

publication_id	term	label	def
IEC 62390, ed. 1.0 (2005-01)	resource logical device	3.1.23	[SOURCE XXXXX]
IEC 60987, ed. 2.0 (2007-08)	fault tree analysis FTA	3.6	[SOURCE XXXXX]

publication_id	term	label	def
IEC 61158-6-10, ed. 3.0 (2014-08)	message frame	3.2.68	[SOURCE XXXXX]
IEC 62065, ed. 2.0 (2014-02)	alarm high-priority alert	3.1.2	[SOURCE XXXXX]
IEC 61158-3-13, ed. 2.0 (2014-08)	application process application layer task	3.3.1	[SOURCE XXXXX]

5.5.19 Definition Addresses Use But Not Concept

In these examples, the definitiens does not clarify, or only partly clarifies, the term (as a concept), but instead describes its usage. Typical trigger words/structures are, for example, *acronym, denotes*, *designation*, *expression, used for describing.*

Examples:

Table 25 – Examples of definitions addressing the use but not the concept

publication_id	term	label	def
IEC 60068-2-81:2003	damping	3.4	generic term ascribed to the numerous energy dissipation mechanisms in a system. In practice, damping depends on many parameters, such as the structural system, mode of vibration, strain, applied forces, velocity, materials, joint slippage, etc.
IEC 60393-1:2008	grade	2.2.3	term indicating additional general characteristics concerning the intended application, for example, long-life applications which may only be used in combination with one or more words (for example, long-life grade) and not by a single letter or number. Figures to be added after the term “grade” should be Arabic numerals
IEC 60744:2018	dependability	3.3	general term describing the overall trustworthiness of a system; i.e. the extent to which reliance can justifiably be placed on this system. Reliability, availability, and safety are attributes of dependability
IEC 60913:2013	drag factor	3.6.1	It is used to consider the shape of a object exposed to wind. The wind pressure is multiplied by this factor to determine the wind action.
IEC PAS 62883:2014	I/O device	3.1.6	an abbreviation for input and / or output device. A device that provides an input and / or output channel for facilitating explicit interaction between a smart environment and its human users. Input devices, such as a microphone, a keyboard, or a mouse, can capture an instruction or response that is provided by a human user and represent it in terms of data in the virtual realm (Figure 6). Upon receive of data within the virtual realm that is intended to be presented to human users, output devices, such as displays and loudspeakers, can make it perceivable to the addressed humans |FIGURE|

No corrected examples are offered for this category.

5.6 Example STS Output

STS is a software tool based on a machine-learning engine (MLE) which attempts to gather the varying definiens of a homonym into semantic classes according to similarity. There are three broad features which STS remarks, as listed in Section 1, namely:

- minor syntactic differences,
- minor semantic differences,
- larger differences.

As an example of minor syntactic differences, consider the term *ciphertext* defined in IEC 62443-1-1:2009 and IEC 62443-3-1:2009:

- *ciphertext* has as definiens in IEC 62443-1-1:2009 *data that has been transformed by encryption so that its semantic information content (i.e., its meaning) is no longer intelligible or directly available,*
- *ciphertext* has as definiens in IEC 62443-3-1:2009 *data that have been transformed by encryption so that the semantic information content (i.e., meaning) is no longer intelligible or directly available.*

The differences between these two definitions are (with alternatives separated by "/") *has/have; its semantic…/the semantic…; its meaning/meaning*. These differences are essentially purely syntactic. It follows the harmonisation task here is simply to pick one of each of the three alternatives.

This may also be seen in Definitions 1-4 of the term *risk* in Figure 1. Definition 5 differs from Definitions 1-4 largely in the inclusion of a field of use <*safety*>, which is not semantically part of the definiens but rather a meta-characteristic.

The box below shows a number of definitions of the term "*risk*" occurring in standards whose terminology is included in the IEC Glossary. STS grouped them into four semantic classes, indicated by the horizontal lines: Definitions 1-12 are similar; Definitions 13-15 are similar; Definitions 17-18 are similar, and Definition 16 is unique (a semantic class of 1 entry).

Some syntactically-classifiable differences are indicated in Figure 1 in bold font, to allow easier comparison of the definitions. Visual inspection indicates that Definitions 1-4 differ in the existence/absence of a determiner; Definition 5 includes a field of use designator. These are syntactic differences.

Definitions 6 and 7 add a parenthetical clause. This is a minor semantic difference. It is semantic because the definitions include an additional meaningful phrase, namely that in parentheses.

Definitions 8 and 9 include the terms *probable* (associated with *probability*), *occurrence*, *harm* and *severity*, as key terms in common with Definitions 1-7. The terms *harm* and *severity* are themselves explicitly defined in applicable standards, *occurrence* and *probable* / *probability* as general universally-meaningful terms are not.

Definition 10 includes key terms *likelihood*, associated with *probability*; *combined*, associated with *combination*, and the defined term *harm*.

Definitions 11 and 12 are obviously similar. Definition 11 includes the key terms *combination occurrence*, as well as the defined terms *harm* and *severity*. Definition 12 includes in addition *probability* as well as the defined term *hazardous event*, whose definiens contains the defined term *harm*.

It is to be noted that the reasons given above are "manual", human reasons why STS might have grouped the definitions as it did. STS is driven by an MLE, so by virtue of the technology it is not amenable to such a rationale to explain why STS did what it did. However, it is clear that any halfway adequate MLE will, for example, group into the same semantic class those definitions which only differ by a determiner.

Table 26 – 18 varying definitions for the term "risk"

1.	**a** combination of the probability of occurrence of harm and the severity of that harm
2.	**a** combination of the probability of **the** occurrence of harm and the severity of that harm
3.	combination of the probability of occurrence of harm and the severity of that harm

4. combination of the probability of **the** occurrence of harm and the severity of that harm
5. **<safety>** combination of the probability of occurrence of harm and the severity of that harm
6. combination of the probability of occurrence of harm **(i.e., physical injury or damage to health**) and the severity of that harm
7. combination of the probability of occurrence of harm (i.e., physical injury or damage to health) and the severity of that harm
8. probable rate of occurrence of a hazard causing harm and the degree of severity of harm
9. probable rate of occurrence of a hazard causing harm and the degree of severity of **the** harm
10. the likelihood, combined with the effect, of loss damage or harm
11. combination of the rate of occurrence of accidents and incidents resulting in harm (caused by a hazard) and the degree of severity of that harm
12. the combination of the frequency, or probability, and the consequence of a specified hazardous event

13. **<security>** expectation of loss expressed as the probability that particular threat will exploit a particular vulnerability with a particular consequence
14. expectation of loss expressed as the **likelihood** that a particular threat will exploit a particular vulnerability with a particular consequence
15. expectation of loss expressed as the **probability** that a particular threat will exploit a particular vulnerability with a particular consequence

16. probability of a failure due to a tin whisker

17. measure of the potential inability to achieve overall program objectives within defined cost, schedule, and technical constraints
18. effect of uncertainty on objectives

Bibliography

[1] Derek Goldrei, Propositional and Predicate Calculus: A Model of Argument, Springer-Verlag London Limited, 2005.

[2] Leslie Lamport, Specifying Systems, Addison-Wesley, 2003.

VDE Verband der Elektrotechnik
Elektronik Informationstechnik e.V.

Merianstraße 28
63069 Offenbach am Main
Tel. +49 69 6308-0
service@vde.com
www.vde.com

VDE