| Value | Description | Example | Scope | Notes |
| --- | --- | --- | --- | --- |
| terminology | An incorrect term or a term from the wrong domain was used or terms are used inconsistently. | * The localization had “Pen Drive” when corporate terminology specified that “USB Stick” was to be used.
* The localized text inconsistently used "Start" and "Begin".
* A text renders the Hungarian term recsegőhid as “buzzer bridge” in English (a literal translation), but the term used in English should be “wedge block,” as specified in a terminology list supplied to the translator.
 | S or T | This value must not be used for simple typographical errors or word choice not related to defined terminologies. For example, a mistyping of “pin” as “pen” or the use of “imply” instead of “infer” (mistaking two commonly confused words) would not count as terminology issues and should be categorized as either spelling errors or mistranslations, depending on the nature of the issue. Terminology refers only to cases where incorrect choices about terms (either formal or commonly defined in a domain) are involved. |
| mistranslation | The content of the target mistranslates the content of the source. | * The English source reads "An ape succeeded in grasping a banana lying outside its cage with the help of a stick" but the Italian translation reads "l'ape riuscì a prendere la banana posta tuori dall sua gabbia aiutandosi con un bastone" ("A bee succeeded...")
 | T | Issues related to translation of specific terms related to the domain or task-specific language should be categorized as terminology issues. |
| omission | Necessary text has been omitted from the localization or source. | * One or more segments found in the source that should have been translated are missing in the target.
* After an alignment, a verification tool flags the pairs of aligned segments where the target has no corresponding source because of incorrect segmentation or some alignment issue. In such case the 'omission' type may apply to the source entry.
 | S or T | This value should not be used for missing whitespace or formatting codes, but instead should be reserved for linguistic content. |
| untranslated | Content that should have been translated was left untranslated. | * The source segment reads "The Professor said to Smith that he would hear from his lawyer" but the Hungarian localization reads "A professzor azt mondta Smithnek, hogy he would hear from his lawyer."
 | T | omission takes precedence over untranslated. Omissions are distinct in that they address cases where text is not present, while untranslated addresses cases where text has been carried from the source untranslated. |
| addition | The translated text contains inappropriate additions. | * The translated text contains a note from the translator to himself to look up a term; the note should have been deleted but was not.
 | T |  |
| duplication | Content has been duplicated improperly. | * A section of the target text was inadvertently copied twice in a copy and paste operation.
 | T |  |
| inconsistency | The text is inconsistent with itself or is translated inconsistently (NB: not for use with terminology inconsistency). | * The text states that an event happened in 1912 in one location but in another states that it happened in 1812.
* The translated text uses different wording for multiple instances of a single regulatory notice that occurs in multiple locations in a series of manuals.
 | S or T |  |
| grammar | The text contains a grammatical error (including errors of syntax and morphology). | * The text reads "The guidelines says that users should use a static grounding strap."
 | S or T |  |
| legal | The text is legally problematic (e.g., it is specific to the wrong legal system). | * The localized text is intended for use in Thailand but includes U.S. regulatory notices.
* A text translated into German contains comparative advertising claims that are not allowed by German law.
 | S or T |  |
| register | The text is written in the wrong linguistic register of uses slang or other language variants inappropriate to the text. | * A financial text in U.S. English refers to dollars as "bucks".
 | S or T |  |
| locale-specific-content | The localization contains content that does not apply to the locale for which it was prepared. | * A text translated for the Japanese market contains call center numbers in Texas and refers to special offers available only in the U.S.
 | S or T | Legally inappropriate material should be classified as legal. |
| locale-violation | Text violates norms for the intended locale. | * A text localized into German has dates in YYYY-MM-DD format instead of in DD.MM.YYYY.
* A text for the Irish market uses American-style foot and inch measurements instead of centimeters.
* A text intended for a U.S.-based audience uses U.K. spellings such as “centre” and “colour.”
 | S or T | This value should be used for spelling errors only if they relate specifically to locale expectations (e.g., a text consistently uses British instead of U.S. spellings for a text intended for the U.S.). If these errors are not systematic (e.g., a text uses U.S. spellings but has a single instance of “centre”), they should instead be counted as spelling errors. |
| style | The text contains stylistic errors. | * Company style guidelines dictates that all individuals be referred to as Mr. or Ms. with a family name, but the text refers to “Jack Smith”.
 | S or T |  |
| characters | The text contains characters that are garbled or incorrect or that are not used in the language in which the content appears. | * A text should have a '•' but instead has a '¥' sign.
* A German text erroneously uses û, ô, and â instead of the appropriate 'ü', 'ö', and 'ä'. A Japanese text has been garbled and appears with Devanagari characters.
 | S or T | Characters should be used in cases of garbling or systematic use of inappropriate characters, not for spelling issues where individual characters are replaced with incorrect one. |
| misspelling | The text contains a misspelling. | * A German text misspells the word "Zustellung" as "Zustlelung".
 | S or T |  |
| typographical | The text has typographical errors such as omitted/incorrect punctuation, incorrect capitalization, etc. | * An English text has the following sentence: "The man whom, we saw, was in the Military and carried it's insignias".
 | S or T |  |
| formatting | The text is formatted incorrectly. | * Warnings in the text are supposed to be set in italic face, but instead appear in bold face.
* Margins of the text are narrower than specified.
 | S or T |  |
| inconsistent-entities | The source and target text contain different named entities (dates, times, place names, individual names, etc.) | * The name "Thaddeus Cahill" appears in an English source but is rendered as "Tamaš Cahill" in the Czech version.
* The date "February 9, 2007" appears in the source but the translated text has "2. September 2007".
 | S or T |  |
| numbers | Numbers are inconsistent between source and target. | * A source text states that an object is 120 cm long, but the target text says it is 129 cm. long.
 | S or T | Some tools may correct for differences in units of measurement to reduce false positives (e.g., a tool might adjust for differences in values between inches and centimeters to avoid flagging numbers that seem to be different but are in fact equivalent). |
| markup | There is an issue related to markup or a mismatch in markup between source and target. | * The source segment has five markup tags but the target has only two.
* An opening tag in the text is missing a closing tag.
 | S or T |  |
| pattern-problem | The text fails to match a pattern that defines allowable content (or matches one that defines non-allowable content). | * The tool disallows the regular expression pattern ['"”’][\.,] but the translated text contains "A leading “expert”, a political hack, claimed otherwise."
* A tool uses a regular expression to ensure that the content of an element is an IRI and flags what appears to be a malformed IRI.
 | S or T | Defining what is or is not an allowable pattern is up to the processing application and is beyond the scope of this specification. Best practice would be to use the Comment attribute to specify the pattern that led to the issue. |
| whitespace | There is a mismatch in whitespace between source and target content or the text violates specific rules related to the use of whitespace. | * A source segment starts with six space characters but the corresponding target segment has two non-breaking spaces at the start.
* The text uses a run of 12 space characters instead of a tab character to align numbers in a table.
* Two space characters appear after a period even though only a single period should be used.
 | S or T |  |
| internationalization | There is an issue related to the internationalization of content. | * A line of programming code has embedded language-specific strings.
* A user interface element leaves no room for text expansion.
* A form allows only for U.S.-style postal addresses and expects five digit U.S. ZIP codes.
 | S or T | There are many kinds of internationalization issues. This value is therefore very heterogeneous in what it can refer to. |
| length | There is a significant difference in source and target length. | * The translation of a segment is five times as long as the source.
 | S or T | What constitutes a "significant" difference in length is determined by the model referred to in the locQualityIssueProfileRef. |
| non-conformance | The content is deemed to show poor statistical conformance to a reference corpus. Higher severity values reflect poorer conformance. | The sentence "The harbour connected which to printer is busy or configared not properly." would have poor conformance. | S or T | Non-conformance is determined through the use of multiple statistical measures of similarity to a corpus of known-good content. For example, in a system that uses classification techniques the poor conformance might be a function of combined incorrect terminology, wrong spelling and bad grammar, or other features as determined by the system. |
| uncategorized | The issue either has not been categorized or cannot be categorized. | * A new version of a tool returns information on an issue that has not been previously checked and that is not yet classified.
* A text is defective in ways that defy categorization, such as the appearance of nonsense garbled text of unknown origin (e.g., a translation shows an unintelligible result and/or appears unrelated to the source material).
 | S or T | This value has the following uses: 1. A tool can use it to pass through quality data from another tool in cases where the issues from the other tool are not classified (for example, a localization quality assurance tool interfaces with a third-party grammar checker).
2. A tool’s issues are not yet assigned to values, and, until an updated assignment is made, they may be listed as uncategorized. In this case it is recommended that issues be assigned to appropriate values as soon as possible since uncategorized does not foster interoperability.
3. uncategorized can be used where a portion of text is defective in a way that defies assignment to a value in either the originating system or in any other ITS localization quality markup to indicate that it is uncategorizable.
 |
| other | Any issue that cannot be assigned to any values listed above. |  | S or T | * This value allows for the inclusion of any issues not included in the previously listed values. This value must not be used for any tool- or model-specific issues that can be mapped to the values listed above.
* In addition, this value is not synonymous with uncategorized in that uncategorized issues may be assigned to another precise value, while other issues cannot.
* If a system has an "miscellaneous" or "other" value, it must be mapped to this value even if the specific instance of the issue might be mapped to another value.
 |

Note: uncategorized is used for issues that have not (yet) been categorized into a more specific value. For example, an automatic process might flag issues for attention but not provide any further detail or categorization: such issues would be listed as “uncategorized” in ITS 2.0. It may also be sued when the exact nature of an issue is unclear and it cannot be categorized as a result (e.g., text is seriously garbled and the cause it unclear). By contrast other is used when the nature of an issue is clear but it *cannot* be categorized in one of the ITS 2.0 categories (or when a model or tool has its own “other” category). For example, in translation of subtitles there is a “respeaking” error category that does not correspond to any ITS 2.0 category; respeaking errors would therefore be categorized as other in ITS 2.0.

**Which of the following localization quality issues are easy to understand?**

|  |  |  |  |
| --- | --- | --- | --- |
| **Id** | **Category**  | **Votes**  | **Comment**  |
| 21 | pattern-problem  |  |  |
| 1 | terminology  | 8  |  |
| 11 | local-specific-content  | 5  |  |
| 8 | grammar  | 5  | I would also regard this as something like 'linguistics'.  |
| 6 | duplication  | 5  | * I would also regard this as "mistranslation".
* We just consider this an addition (to keep things simple).
 |
| 5 | addition  | 5  |  |
| 4 | untranslated  | 5  | * I would also simply regard this as "mistranslation" (or "omission").
* I think it is good to have these detailed categories to be able to classify the errors more accurately.
 |
| 3 | omission  | 5  |  |
| 23 | internationalization  | 4  |  |
| 17 | formatting  | 4  | I would add here "and layout". I would also include "characters" here.  |
| 16 | typographical  | 4  | * I would also regard this as something like 'linguistics'.
* Style?
 |
| 15 | misspelling  | 4  | I would also regard this as something like 'linguistics'.  |
| 14 | characters  | 4  | * I would also regard this as "formatting".
* Issues relating to characters have usually been due to a technical issue and were not always visible to the translator. Would we still count this as an error by the translator?
 |
| 12 | locale-violation  | 4  | I would also regard this as "local-specific content". Translators need localization guidelines.  |
| 7 | inconsistency  | 4  | For the first example given in the document, I would just class the inconsistency of the date as a typographical error usually. I agree with the inconsistency of texts that should be standard.  |
| 2 | mistranslation  | 4  |  |
| 26 | uncategorized  | 3  |  |
| 25 | non-conformance  | 3  |  |
| 24 | length  | 3  |  |
| 22 | whitespace  | 3  |  |
| 20 | markup  | 3  |  |
| 19 | numbers  | 3  |  |
| 13 | style  | 3  | I think this is a tricky one because S&G for writing may for example not provide guidance for all documentation types.  |
| 10 | register  | 3  | * I would also regard this as "style".
* Agree
 |
| 9 | legal  | 3  | * I would also regard this as "local-specific content". Translators need localization guidelines.
* Agree
 |
| 27 | other  | 2  |  |
| 18 | inconsistent-entities  | 1  |  |

**Which of the following localization quality issues cannot be understood?**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Category**  | **Votes**  | **Comment**  |
| 27 | other  |  |  |
| 26 | uncategorized  |  |  |
| 17 | formatting  |  |  |
| 16 | typographical  |  |  |
| 15 | misspelling  |  |  |
| 14 | characters  |  |  |
| 12 | locale-violation  |  |  |
| 11 | locale-specific-content  |  |  |
| 9 | legal  |  |  |
| 8 | grammar  |  |  |
| 6 | duplication  |  |  |
| 4 | untranslated  |  |  |
| 2 | mistranslation  |  |  |
| 1 | terminology  |  |  |
| 21 | pattern-problem  | 4  |  |
| 18 | inconsistent-entities  | 3  | * I think especially the date example is a "mistranslation".
* I think the examples point more in the direction of "locale-violation"
* Many things such as names, dates should be changed to be appropriate for the target language audience. I found the explanation here confusing - seemed to indicate it is wrong to change the name, date format, etc.
 |
| 25 | non-conformance  | 2  |  I would regard this as "mistranslation".  |
| 22 | whitespace  | 2  |  I would include this in "formatting and layout".  |
| 20 | markup  | 2  | I would regard this as an "omission" (including formatting omissions).  |
| 19 | numbers  | 2  | * I think the example refers to a "mistranslation".
* Agree with Marja. In example, number is not incorrect, unit of measurement is, so perhaps this should be named unit of measurement instead.
 |
| 10 | register  | 2  |  |
| 24 | length  | 1  | If the translation is correct and readable, and is not cut off in the user interface, for example, I don't think it's an error. But I can't think of any correct or wrong examples here!  |
| 23 | internationalization  | 1  | I understand this as a source (text) issue; if for example a form only allows U.S. postal addresses, it's not a translation error.  |
| 13 | style  | 1  | I understand "style" is a wider area. It could also mean correct industry-specific language, adherence to appropriate style guide(s), using the correct registry, maybe also readability.  |
| 7 | inconsistency  | 1  | Are we talking about source text errors or translation errors? If the source texts gives two different dates for actually the same year (see example), this is not a translation error.  |
| 5 | addition  | 1  |  I would regard this as "mistranslation".  |
| 3 | omission  | 1  |  Why shouldn't this value be used for missing formatting codes?  |

linguistic

 terminology, inconsistency, grammar, register, misspelling, typographical

content

 legal, locale-specific-content, locale-violation, style, inconsistent-entities, internationalization, non-conformance

technical

 characters, formatting, pattern-problem

equivalence

 mistranslation, omission, untranslated, addition, duplication, numbers, markup, whitespace

other

 length, uncategorized, other

Some fall into two categories (e.g. whitespace)