

# Quality assurance

- Version und Versionierung
  - Types of quality assurance in translate5
- Overview
- QA statistics
  - False positives
    - Marking reported errors as false positives individually
    - Marking entire error categories as false positive
  - Manual QA (complete segment)
  - Manual QA (inside segment)
  - Status
- Error categories
  - Consistency
  - Internal tags
  - LanguageTool
  - Empty segments
  - White Space at the beginning/end
  - Length check
  - Manual QA (complete segment)
  - Manual QA (inside segment)
  - Usage of TM matches
  - Terminology
  - Numbers

## Version und Versionierung

Includes functions of the application up to version	7.2.4
Current translate5 version	7.3.2

Version	Published	Changed By	Comment
<b>CURRENT (v. 4)</b>	Apr 11, 2024 12:17	 Marion Gubler	updated to versio
v. 3	Feb 16, 2024 16:05	 Marion Gubler	
v. 2	Feb 12, 2024 16:15	 Marion Gubler	

[Go to Page History](#)

In the editor, you can mark text or complete segments for quality management purposes. For example, you can mark segments with stylistic or grammatical errors and then carry out statistical analyses using this data. You can assign the error categories in the "Quality assurance" section of the right-hand editor panel.

## Types of quality assurance in translate5

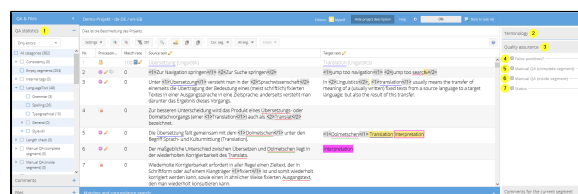
translate5 distinguishes between **automated quality assurance** and **manual quality assurance**:

- During manual quality assurance, a user manually assigns an error category to a complete segment or part of it.
  - Manual QA complete segment: Here, segments can be marked with various individually configured error categories.
  - Manual QA in the segment: Here, parts of a segment can be commented on and assigned an error category in accordance with the [Multidimensional Quality Metrics \(MQM\) framework](#) or with individual error categories.
- Automatic quality assurance assigns an error category and/or comment to a complete segment or part of it.
  - Automated QA complete segment covers cases such as consistency (equal source segments translated differently) or empty segments, length check or (missing) processing of 100% or fuzzy matches.
  - Automated QA in the segment covers cases such as numbers differing between source and target segment, incorrect thousands separators and errors found by the spell checker.

## Overview

Four editor areas are available for quality assurance:

1. The left-hand editor panel "**QA statistics**", where you can filter by specific manually or automatically assigned QS categories.
2. The "**Terminology**" section of the right-hand editor panel: Terminology recognized by



the TermCollection is displayed here.

3. The "Quality assurance" section of the right-hand editor panel, where you can set and control the QA categories at segment and content level that can be filtered in the left-hand editor panel. This section is subdivided into:
  4. "False positives": You can mark or unmark errors found by the quality assurance as false positives.
  5. "Manual QA (complete segment)": Here, quality problems at segment level can be pointed out using, various predefined categories.
  6. "Manual QA (inside segment)": Here, quality problems within segments can be marked manually by means of comments and severity, or using the predefined QA categories.
  7. "Status": Here, a status can be set per segment, e.g. stating that a segment needs to be checked again.




---

## QA statistics

In this section of the editor, you can see the list of error categories that are checked during quality assurance and either assigned automatically or manually at segment level. You can filter and display the segments based on these categories by ticking the categories in the filter field. If you select one of the main categories, all subcategories of this category will also be selected. The same applies to deselecting a main category. The very first line "All categories" can be used to select or deselect all main categories and their subcategories. Error categories that do not currently apply to any of the segments cannot be selected. Using the drop-down above the error categories, you can choose whether you want to:

- Show all
- Only errors
- Only False positives

. You can combine the drop-down selection with the selected categories for filtering.

	False positives are bits of text that were flagged as false by the automated quality assurance, but are in fact not errors. This happens, for example, when there is a number in the source segment that is translated using the corresponding word. This segment would be marked with a numbers error, because the system cannot find the number in the target segment, and would therefore be a false positive.
	Error categories whose counter is a "0" as there are no segments to which this category applies cannot be selected.
	The categories that refer to errors that block the export of the task, such as structural tag errors, are marked with red font colour.

---

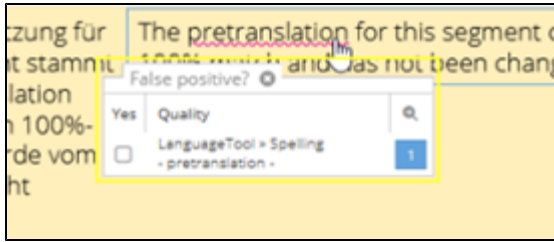
## False positives

This section of the right-hand editor panel shows all error categories that have been assigned to the currently open segment. The individual error categories can be set to "False positive" here, which means that this error category was applied incorrectly.

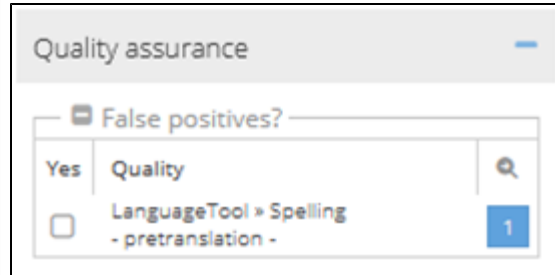
### Marking reported errors as false positives individually

Directly in the segment	In the right-hand editor panel
-------------------------	--------------------------------

Right-click on the relevant part of the segment. A small pop-up appears where you can place a tick in the “Yes” column. The reported error is now marked as a false positive.



In the “Quality assurance” section of the right-hand editor panel, the error categories assigned to the currently open segment are displayed. Place a tick in the “Yes” column of the list for categories that you want to mark as false positive.



Please note that setting the “false positive” tick will be saved immediately, even if editing of the segment is cancelled without saving.

## Marking entire error categories as false positive

You can set complete error categories reported by LanguageTool and terminology recognition to false positive.

Directly in the segment	In the right-hand editor panel
Instead of ticking the box on the left-hand side of the false positive menu, click on the box with the number on the right-hand side.	First, tick the “Yes” box of the relevant category in the “Quality assurance” area of the right-hand editor panel. Then click on the blue number that appears in the right-hand column, which indicates the number of errors in this error category. The number will no longer appear with blue highlighting. This can be undone by clicking on the number again.



Please note that setting the “false positive” tick will be saved immediately, even if editing of the segment is cancelled without saving.

## Manual QA (complete segment)

Here, you can assign error categories on segment level.

	Please note that the assignment of error categories is saved immediately, even if editing of the segment is cancelled without saving.
	These error categories can be defined and adjusted individually for each task in the system configuration.

## Manual QA (inside segment)

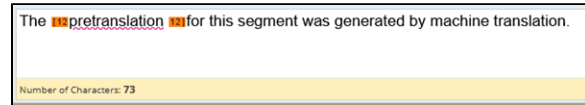
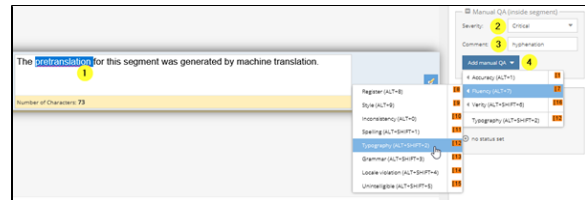
During manual quality assurance in the segment, a user can assign an error category for one or several words within a segment. To do so, proceed as follows:

1. Mark the relevant bit of text in the segment.
2. Select the severity of the error from the drop-down (optional).
3. Add a comment (optional).
4. Select an error category from the “Add manual QA” drop-down.

The “Add Manual QA” drop-down is multidimensional. If you want to display the submenu of one of the categories, just move the cursor over it. However, you can also directly assign one of the error categories appearing on the first level.

The selected segment passage is then surrounded by a pair of tags with an orange background.

The system remembers most recently used error categories and offers them for direct selection among the categories on the first drop-down level.

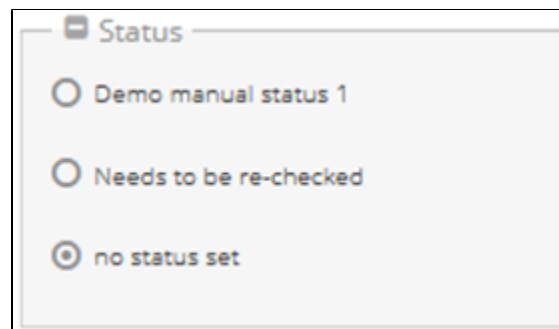


	The categories displayed here can be defined and adjusted individually for each task in the “QM-Subsegment_Issues.xml” file in the import zip.
	If no text is marked before selecting the error category, no category will be assigned.
	The error category assignments inside the segment are retained only when the segment is saved.
	For more efficiency, the error categories can also be assigned via a keyboard shortcuts, see <a href="#">Keyboard shortcuts in the editor</a> .

## Status

In this area, a manual quality assurance status can be assigned to segments. A status can, e.g. indicate to the reviser or second proofreader which segments should be given special attention to.

	The statuses displayed here can be defined in the configuration of your translate5 system.
--	--



## Error categories

This section gives an overview of the error categories.

### Consistency

<b>Inconsistent source</b>	The target segment contains the same translation that was used for a deviating source segment.
<b>Inconsistent target</b>	The target segment has been translated differently from another target segment whose source segment content matches the current one.

### Internal tags

<b>Internal tags have an incorrect structure</b>	<p>The order of the tags in the target segment is incorrect in one of the following ways</p> <ul style="list-style-type: none"> <li>e.g. a closing tag has been inserted before the opening tag:  <div>The term <b>&lt;/1&gt;</b>mediation<b>&lt;1&gt;</b> is often used in language didactics.</div> </li> <li>or the tags were inserted incorrectly nested or are overlapping:  <div>(See literary translation<b>&lt;/1&gt;</b>) At the time of Romanticism, many intellectuals were also theoretically concerned with translating, like <b>&lt;2&gt;</b>Johann Wolfgang von Goethe<b>&lt;/2&gt; &lt;1&gt;</b>, <b>&lt;3&gt;</b>Friedrich Schleiermacher or<b>&lt;4&gt;</b><b>&lt;/3&gt;</b>Wilhelm von Humboldt<b>&lt;/4&gt;</b> . </div> </li> </ul>
<b>Internal tags have been added</b>	The target segment contains tags that are not present in the corresponding source segment.
<b>Internal tags missing</b>	Not all tags from the source segment are present in the target segment.
<b>Whitespace has been added</b>	The target segment contains at least one additional space, tab, or line break before/after a tag compared to the source segment. It makes no difference whether the space is missing before or after the tag.
<b>Whitespace has been removed</b>	The target segment contains at least one space, tab, or line break less before/after a tag compared to the source segment. It makes no difference whether the space is missing before or after the tag.

## LanguageTool

See also [Terminology, style and spell checking](#).

<b>Grammar</b>	The text contains a grammatical error (including errors in syntax and morphology).
<b>Spelling</b>	The text contains a spelling mistake.
<b>Typographical</b>	The text has typographical errors, such as missing/incorrect punctuation, incorrect capitalization, etc.
<b>General &gt; Characters</b>	The target segment contains symbols that are wrong, in the wrong order, or that are not used in the target language.
<b>General &gt; Duplication</b>	The target segment contains at least one word twice.
<b>General &gt; Inconsistency</b>	The text is inconsistent in that it uses different translations for the same term in different places (not coupled with terminology recognition).
<b>General &gt; Legal</b>	The text is legally problematic (e.g. it refers to the wrong legal system, or uses expressions and phrases that are not common for the target language's legal area).
<b>General &gt; Uncategorized</b>	The identified error is either uncategorized or cannot be categorized (due to incomprehensible strings).
<b>Style &gt; Register</b>	The translation is written in the wrong language register or uses slang or other language variants that are not appropriate for the target text.
<b>Style &gt; Locale-specific content</b>	The translation contains passages that do not apply to the target language culture or country.
<b>Style &gt; Locale-violation</b>	The text violates the standards for the intended locale.
<b>Style &gt; General style</b>	The text contains stylistic errors.
<b>Style &gt; Pattern problem</b>	The text does not match a pattern that defines acceptable content (or matches a pattern that defines non-acceptable content).

<b>Style &gt; Whitespace</b>	The kind or number of white spaces in the source and target segment do not match, or the target segment violates certain rules for the use of white space.
<b>Style &gt; Terminology</b>	A wrong term, or a term from the wrong area was used, or terminology was used inconsistently.
<b>Style &gt; Internationalization</b>	There is a problem related to the internationalization of content.

## Empty segments

<b>Empty segments</b>	There are segments that have not been translated yet.
-----------------------	---

## White Space at the beginning/end

<b>Line break at the beginning</b>	The segment contains an unnecessary line break right at the beginning.
<b>Line break at the end</b>	The segment contains an unnecessary line break at the end.
<b>Space after line break</b>	The segment contains an unnecessary space after the line break.
<b>Non-breaking space at the beginning</b>	The segment contains a non-breaking space right at the beginning.
<b>Non-breaking space at the end</b>	The segment contains an unnecessary non-breaking space at the end.
<b>Space before line break</b>	The segment contains an unnecessary space before the line break.
<b>Segment ends with a space followed by a tag</b>	The segment contains a tag at the end that is preceded by an unnecessary space.
<b>Tab at the beginning</b>	The segment contains an unnecessary tab stop right at the beginning.
<b>Tab at the end</b>	The segment contains an unnecessary tab stop at the end.
<b>Segment begins with a tag followed by a space</b>	The segment contains a tag at the beginning that is followed by an unnecessary space.

## Length check

<b>Segment relevantly shorter than allowed (more than 20% shorter or at least 100 Pixel shorter or at least 20 characters shorter)</b>	Indicates that the segment's character count is significantly below layout-dependent length specification limits.
<b>Segment longer than allowed</b>	The segment content exceeds the permissible length specification.
<b>Segment has too many lines</b>	The segment contains more line breaks than permitted (e.g. relevant for subtitle files).
<b>Segment shorter than allowed.</b>	The segment content falls short of the permissible length specification.

## Manual QA (complete segment)

The categories displayed here can be configured individually. The following categories are available by default:

- Demo QM error 1
- Mistranslation
- Terminology problem
- Fluency problem
- Inconsistency

## Manual QA (inside segment)

These default quality categories follow the [Multidimensional Quality Metric \(MQM\)](#) model that is based on the LISA QA model, but structured modally, as developed and defined by QTLaunchpad. QTLaunchpad was an EU supported project by the [German Research Center for Artificial Intelligence \(DFKI\)](#), [Dublin City University](#), [University of Sheffield](#) and the [Athena Institute for Language and Speech Processing](#). The categories are as follows:

- Accuracy
  - Mistranslation
    - Terminology
  - Omission

- Addition
  - Untranslated
- Fluency
  - Register
  - Style
  - Inconsistency
  - Spelling
  - Typography
  - Grammar
  - Locale violation
  - Unintelligible
- Verity
  - Completeness
  - Legal requirements
  - Local applicability



The categories displayed here can be defined and adjusted individually for each task in the “QM-Subsegment\_Issues.xml” file in the import zip.

## Usage of TM matches

<b>Edited 100% match</b>	Indicates that a 100% match from the translation memory has been edited.
<b>Unedited fuzzy match</b>	Indicates that a fuzzy match from the translation memory has not been edited.

## Terminology

<b>Forbidden in source</b>	Indicates that a term has been used that is flagged as forbidden for the source language.
<b>Forbidden in target</b>	Indicates that a term has been used that is flagged as forbidden for the target language.
<b>Not defined in target terminology</b>	Indicates that the segment contains a term that is only defined in the source language, but not yet in the target language of the corresponding TermCollection.
<b>Not found in target</b>	Indicates that there is a match with the TermCollection in the source segment, but not in the target segment.

## Numbers

<b>Numbers source target</b>	Indicates that the numbers contained in the target segment do not match those in the source segment.
<b>Format alteration (ordinal numbers, leading zero etc.)</b>	Indicates that the notation or format of the numbering/ordinals in the target segment differs from the one used in the source segment.
<b>Different character/formatting for number interval</b>	Indicates that the notation or format of a number interval in the target segment differs from the one used in the source segment.
<b>1000-separator not allowed</b>	Indicates that a thousands separator that is not permitted for the target language has been used in the target segment.
<b>Dubious number from source unchanged in target</b>	Indicates that a possibly incorrect number was detected in the source segment that was transferred to the target segment and not edited.
<b>Alphanumeric character sequences: Differences</b>	Indicates differences between the source and target segments regarding alphanumeric strings, e.g. article names that have not been transferred correctly.
<b>Format alteration (date information etc.)</b>	Indicates that the notation or format of dates in the target segment is not the same as in the source segment.
<b>Separator(s) not localised</b>	Indicates that the thousands separator was transferred from the source segment to the target segment and not edited to match the target language's requirements.
<b>Formatting of 1000s number altered</b>	Indicates that the notation or format of the thousands numbers in the target segment differs from the one in the source segment.
<b>Numerals from source found as number in target</b>	Indicates that the source segment contains a numeral (one, two, three, etc.), for which a number (1, 2, 3, etc.) was used in the target segment.

<b>Number from source found as numeral in target.</b>	Indicates that the source segment contains a number (1, 2, 3, etc.), for which a numeral (one, two, three, etc.) was used in the target segment.
---	--