INTERNATIONAL STANDARD

ISO 18587

First edition
2017-04

Translation services — Post-editing of machine translation output — Requirements

Services de traduction — Post-édition d'un texte résultant d'une traduction automatique — Exigences

Reference number
ISO 18587:2017(E)

© ISO 2017



COPYRIGHT PROTECTED DOCUMENT

© ISO 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

Contents

Page

Foreword ... **iv**
Introduction ... **v**
1 Scope ... **1**
2 Normative References ... **1**
3 Terms and definitions ... **1**
3.1 Concepts related to machine translation ... 1
3.2 Concepts related to language and content ... 2
3.3 Concepts related to people or organizations ... 3
3.4 Concepts related to translation ... 4
4 Post-editing process ... **5**
4.1 General ... 5
4.2 Pre-production processes ... 5
4.3 Production processes ... 6
4.3.1 Objectives in the post-editing process ... 6
4.3.2 Requirements of post-editing MT output ... 6
4.3.3 Post-editor's tasks ... 7
4.4 Post-production processes ... 7
4.4.1 Final verification and delivery ... 7
4.4.2 Feedback ... 7
5 Competences and qualifications of post-editors ... **7**
5.1 Competences ... 7
5.2 Qualifications ... 8
5.3 Professionalism ... 8
6 Requirements of full post-editing ... **8**
Annex A (informative) **Post-editor training** ... **9**
Annex B (informative) **Light post-editing** ... **10**
Annex C (informative) **Pre-editing** ... **11**
Annex D (informative) **Client–TSP agreements and project specifications** ... **12**
Annex E (informative) **Automatic post-editing** ... **14**
Bibliography ... **15**

© ISO 2017 – All rights reserved

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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

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

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

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

This document was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 5, *Translation, interpreting and related technology*.

© ISO 2017 – All rights reserved

Introduction

The use of machine translation (MT) systems to meet the needs of an increasingly demanding translation and localization industry has been gaining ground. Many translation service providers (TSPs) and clients have come to realize that the use of such systems is a viable solution for translating projects that need to be completed within a very tight time frame and/or with a reduced budget. When an MT system is used, clients can have material translated that can otherwise not be translated; translation costs can be decreased and the launch of products on specific markets, as well as the flow of information, can be accelerated. On the other hand, TSPs are able to:

a) improve translation productivity;

b) improve turn-around times;

c) remain competitive in an environment where clients show an increasing demand for using MT in translation.

However, there is no MT system with an output which can be qualified as equal to the output of human translation and, therefore, the final quality of the translation output still depends on human translators and, for this purpose, their competence in post-editing.

The rate at which MT systems are changing renders it impractical to produce an overarching International Standard on these systems, which could stifle innovation or be ignored by the translation technology development industry.

This document therefore restricts its provisions to that part of the process that begins upon the delivery of the MT output and the beginning of the human process that is known as post-editing.

© ISO 2017 – All rights reserved

Translation services — Post-editing of machine translation output — Requirements

1 Scope

This document provides requirements for the process of full, human post-editing of machine translation output and post-editors' competences.

This document is intended to be used by TSPs, their clients, and post-editors.

It is only applicable to content processed by MT systems.

NOTE For translation services in general, see ISO 17100.

2 Normative References

There are no normative references in this document.

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:

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

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

3.1 Concepts related to machine translation

3.1.1
machine translation
MT
automatic *translation* (3.4.2) of *text* (3.2.6) from one natural language to another using a computer application

[SOURCE: ISO 17100:2015, 2.2.2, modified – reference to translation of speech has been deleted as it is not relevant to this document; also "automated has been replaced by "automatic" in order to avoid confusion with translation memory tools]

3.1.2
machine translation output
MT output
result of *machine translation* (3.1.1)

[SOURCE: ISO 17100:2015, 2.2.3, modified – "outcome" has been changed to "result"]

3.1.3
machine translation system
technology used to perform *machine translation* (3.1.1)

© ISO 2017 – All rights reserved

3.1.4
post-edit
edit and correct *machine translation output* (3.1.2)

[SOURCE: ISO 17100:2015, 2.2.4, modified – the note has been deleted]

3.1.5
full post-editing
process of *post-editing* (3.1.4) to obtain a product comparable to a product obtained by *human translation* (3.4.3)

3.1.6
light post-editing
process of *post-editing* (3.1.4) to obtain a merely comprehensible text without any attempt to produce a product comparable to a product obtained by *human translation* (3.4.3)

3.2 Concepts related to language and content

3.2.1
content
information in any form

EXAMPLE Text, audio, video, etc.

3.2.2
source language
language of the *content* (3.2.1) to be *translated* (3.4.1)

3.2.3
source language content
language *content* (3.2.1) to be *translated* (3.4.1)

[SOURCE: ISO 17100:2015, 2.2.3]

3.2.4
target language
language into which *source language content* (3.2.3) is *translated* (3.4.1)

[SOURCE: ISO 17100:2015, 2.3.6]

3.2.5
target language content
language *content* (3.2.1) *translated* (3.4.1) from *source language content* (3.2.3)

[SOURCE: ISO 17100:2015, 2.3.3]

3.2.6
text
content (3.2.1) in written form

[SOURCE: ISO 17100:2015, 2.3.4]

3.2.7
natural language
NL
language with its origin unknown, but continuously developing sometimes in idiosyncratic ways as is used conventionally for human communication

[SOURCE: ISO/TS 24620, 2.12]

© ISO 2017 – All rights reserved

3.2.8
controlled natural language
controlled language
CNL
subset of *natural languages* (3.2.7) whose grammars and dictionaries have been restricted in order to reduce or eliminate both ambiguity and complexity

Note 1 to entry: As a generic, CNL is an uncountable noun that refers to the abstract properties of all controlled natural languages and not to a particular natural language or application for a specific purpose. It is engineered (i.e. constructed) with a view to reducing or eliminating ambiguity and complexity and aims both to make it easier for human readers (particularly non-native users, non-experts and people with limited comprehension) to read a *text* (3.2.6), and to improve the computational processing of a text.

Note 2 to entry: CNL is an engineered (i.e. constructed) language that is based on a particular natural language, but is more restrictive as regards lexicon, syntax, or semantics, while at the same time preserving most of its natural properties. Here, CNL is a countable noun.

[SOURCE: ISO/TS 24620, 2.6]

3.2.9
segment
unit of *text* (3.2.6) produced for a computer application to facilitate translation

Note 1 to entry: A segment can be a sentence, heading or other unit of text, such as phrase, word or a single character.

3.2.10
locale
set of characteristics, information or conventions specific to the linguistic, cultural, technical and geographical conventions of a target audience

3.2.11
language register
variety of language used for a particular purpose or in a particular social or industrial domain

3.3 Concepts related to people or organizations

3.3.1
client
customer
person or organization that commissions a service from a *TSP* (3.3.5) by formal agreement

Note 1 to entry: The client can be the person or organization requesting or purchasing the service and can be external or internal to the *TSP*'s (3.3.5) organization.

[SOURCE: ISO 17100:2015, 2.4.3]

3.3.2
translator
person who *translates* (3.4.1)

[SOURCE: ISO 17100:2015, 2.4.4]

3.3.3
reviser
person who revises *translation output* (3.4.4)

[SOURCE: ISO 17100:2015, 2.4.5, modified – no need to specify that it is against source language content as it is explained in the definition of "revision"]

© ISO 2017 – All rights reserved

3.3.4
post-editor
person who *post-edits* (3.1.4)

3.3.5
translation service provider
TSP
language service provider that delivers *translation services* (3.4.5)

Note 1 to entry: A TSP can be a translation company, a translation agency, a translation organization (profit, non-profit or governmental), a single freelance translator or post-editor, or an in-house translation department.

[SOURCE: of definition: ISO 17100:2015, 2.4.2, modified]

3.3.6
project manager
PM
person who manages specified aspects of a translation or post-editing project and is responsible for the process

[SOURCE: ISO 17100:2015, 2.4.8, modified – "post-editing" added to make it relevant to this document]

3.4 Concepts related to translation

3.4.1
translate
render *source language content* (3.2.3) into *target language content* (3.2.5) in written form

[SOURCE: ISO 17100:2015, 2.1.1]

3.4.2
translation
set of processes to render *source language content* (3.2.3) into *target language content* (3.2.5)

[SOURCE: ISO 17100:2015, 2.1.2, modified – no need to mention that it is in "written form". Also, the note is deleted in order to avoid confusion with the definition of translation output.]

3.4.3
human translation
translation (3.4.2) performed by a *translator* (3.3.2)

3.4.4
translation output
result of *translation* (3.4.2)

3.4.5
translation service
intangible product that is the result of interaction between *client* (3.3.1) and *TSP* (3.3.5)

[SOURCE: ISO 17100:2015, 2.1.6]

3.4.6
translation memory
TM
electronic collection of *source language* (3.2.2) and *target language* (3.2.4) *segment* (3.2.9) pairs

Note 1 to entry: The purpose of a translation memory is to facilitate the retrieval and use of previously translated content.

© ISO 2017 – All rights reserved

3.4.7
revision
bilingual examination of *target language content* (3.2.5) against *source language content* (3.2.3) for its suitability for the agreed purpose

Note 1 to entry: The term "bilingual editing" is sometimes used as a synonym for revision.

[SOURCE: ISO 17100:2015, 2.2.6]

3.4.8
verification
confirmation by the *TSP's* (3.3.5) *project manager* (3.3.6) that specifications have been fulfilled

[SOURCE: ISO 17100:2015, 2.5.1, modified – TSP added for the sake of clarity]

4 Post-editing process

4.1 General

Post-editing is performed on MT output for the purpose of checking its accuracy and comprehensibility, improving the text, making the text more readable, and correcting errors. Post-editing differs from translation as it involves three texts: the source text, the MT output and the final target text.

There are two main levels of post-editing, light and full; their use depends on the purpose of the translation output and the client's requirements. This document deals with full post-editing. Light post-editing is described in Annex B.

The post-editing process can be automatic or human. Automatic post-editing is described in Annex E. Human post-editing refers to the MT output post-editing process performed by a human post-editor.

The post-editing process can take place in an environment where all the text to be processed has been machine translated or in an integrated environment, where MT and TM, along with consistent terminology management, are fully integrated to produce a multi-modal translation work environment.

Post-edited MT output can or cannot be fed into a translation memory (automatically or partially), depending on the organization's processes and its clients' quality assurance requirements.

4.2 Pre-production processes

The TSP shall have a process in place to determine, in consultation with the client when necessary, whether the source language content is suitable for MT and subsequent post-editing or not since combined MT/post-editing efficiency depends on the MT system, language combination, domain and style of the source language content.

The source language content can be pre-edited before machine translation to facilitate machine processing, improve raw MT output quality, and therefore reduce the post-editing workload. Pre-editing is described in Annex C.

The TSP shall finalize an agreement with the client and retain a record of that agreement. If an agreement is made verbally or by telephone, the TSP shall confirm the agreement and its terms in writing (e.g. by letter, fax or email). The agreement, whether contractual or non-contractual, shall include or reference the commercial terms and the project specifications. The agreement may also call for conformity to this document. Annex D contains a list of the items which can be included in the agreement.

Any deviation from the original agreement shall be agreed upon by all parties before any action is taken that deviates from the original agreement, and the agreed amendment shall be recorded and stored with the original agreement documentation.

The TSP shall ensure that requirements for the post-editing project are identified, documented and accessible to the post-editor. The TSP shall also ensure that the relevant specifications (e.g. quality

© ISO 2017 – All rights reserved