



Open Source CAT-Tools

Manuel de la Cruz Recio & Belén Santana López



Cover

MateCat User Manual

Date: 13/12/2023

This manual is dedicated entirely to **MateCat**, the free and open-source computer-assisted translation (CAT) tool. It offers detailed instructions for creating an account, configuring linguistic assets such as translation memories and glossaries, selecting machine-translation engines, and managing your first project from start to finish.

The manual preserves the colour scheme and imagery of the workshop slides to maintain visual consistency.

Credits

Manual authors and facilitators

- **Dr Manuel de la Cruz Recio** – Department of Translation and Interpreting, University of Salamanca (specialist in Interpreting Studies and Translation Technology).
E-mail: manueldelacruzrecio@usal.es
- **Dr Belén Santana López** – Department of Translation and Interpreting, University of Salamanca (specialist in Translation Studies and Literary Translation).
E-mail: bsantana@usal.es

Project affiliation

- **BAQONDE – Learn in Your Language** project (co-funded by the Erasmus + programme of the European Union).
- **University of Salamanca – Campus of International Excellence.**

Table of Contents

1. [Introduction to MateCat](#)
2. [Creating a MateCat Account](#)
3. [Translation Memories in MateCat](#)
4. [Glossaries and Terminology Management](#)
5. [Machine Translation and AI Features](#)
6. [Starting Your First Project](#)
7. [Working in the Editor](#)
8. [Revision and Quality Assurance](#)
9. [Tips and Best Practices](#)
10. [Additional Resources](#)
11. [Acknowledgements](#)

Introduction to MateCat

MateCat is a web-based CAT tool developed by Translated. It is free to use and open source, which means you can leverage its professional features without licence fees. MateCat supports more than 80 file formats and over 200 languages, making it suitable for a wide variety of translation projects. The platform is designed for individual translators, language service providers and organisations that need to manage multilingual content efficiently.

Key characteristics of MateCat include:

- **Web-based accessibility:** No installation required; you only need a modern web browser and an internet connection.
- **Adaptive machine translation:** MateCat integrates *ModernMT Lite*, an adaptive MT engine that learns from your corrections and improves its suggestions over time.
- **Unlimited public translation memory:** MateCat's public TM contains billions of bilingual segments that can be leveraged during translation to find matches.
- **Private translation memories and glossaries:** You can create private assets that remain confidential and are used exclusively in your projects.
- **Collaboration and outsourcing:** MateCat offers options to outsource translations to professional linguists if you prefer not to translate in-house.

Creating a MateCat Account

1. **Access the website:** Navigate to www.matecat.com using your preferred web browser.
2. **Select “Sign Up”:** On the home page, click the **Sign Up** button to open the registration form.
3. **Choose a sign-in method:** You can sign up using single sign-on (Google, LinkedIn, Microsoft or GitHub) for instant access. Alternatively, complete the form with your name, email address and password, then click **Create Account**.
4. **Confirm your email:** MateCat sends a verification link to the email address you provided. Click the link to activate your account and you will be logged in automatically.
5. **Sign in:** Once your account is activated, use the **Sign In** button on the home page to log in. You can also use any of the single sign-on methods you configured during registration.

If you do not receive the verification email, check your spam folder. If the message is not there, contact the MateCat support team for assistance.

Translation Memories in MateCat

While you translate in MateCat, each confirmed segment is stored in a **translation memory (TM)**. These segments can be reused later when similar sentences appear in the same or a different project. In MateCat you have two options for storing translations:

- **Public TM:** Your translations are saved in a generic public memory accessible to all MateCat users. This contributes to the shared corpus and enables matches from other users' work to appear in your projects.
- **Private TM:** You can create a private memory, identified by a unique alphanumeric key known as the *private TM key*. This memory is linked to specific projects or clients and is only accessible to authorised users.

When a sentence identical or similar to an existing TM segment is encountered, MateCat suggests the stored translation. You can accept it in its entirety or adapt it to the new context. Saving translations in a private TM ensures confidentiality and avoids exposing proprietary content to other users.

Glossaries and Terminology Management

MateCat allows you to upload glossaries to guide translators on how to handle specific terminology. A glossary can be prepared as a spreadsheet in **XLSX**, **XLS** or **ODS** format. It may contain multiple language columns and optional context information such as forbidden terms, domain, subdomain and definition. The minimum requirement is that the file contains at least two language columns.

Key aspects of MateCat's glossary system:

- **Multiple languages:** A single glossary can include many languages, and MateCat automatically generates usable combinations (e.g. EN <> ES).
- **Concept-level information:** Optional columns can be used to indicate forbidden terms, domain or subdomain, and provide definitions. Forbidden terms can trigger warnings when used in the translation.
- **Term-level information:** Each locale can have notes and examples of use. MateCat applies advanced matching to recognise different declensions or capitalisation.

During translation, MateCat underlines glossary terms in the source segment and lists them in the **Glossary** tab. You can consult definitions, domain information and notes; add or edit terms; and highlight forbidden words. Terminology QA automatically flags terms not translated according to the glossary or forbidden words in the target. This helps ensure terminological consistency and adherence to client-specific requirements.

Machine Translation and AI Features

MateCat integrates multiple machine translation engines to provide suggestions alongside translation memory matches. The default engine is **ModernMT Lite**, an adaptive system developed by Translated. ModernMT learns from your corrections and tailors its output to your style and terminology. It supports 200 languages and is designed to improve as you work. For language pairs not supported by ModernMT, MateCat falls back on Google Translate.

In the project settings you can select from other available MT engines (e.g. Lara or a full version of ModernMT) and configure the **MT threshold**. This threshold defines the similarity percentage below which machine translation suggestions are preferred over translation memory matches. The default value is 85 % and can be adjusted before creating a project.

The adaptive MT and threshold settings allow you to balance reuse of your own translations with automated suggestions, enhancing productivity without compromising quality.

Starting Your First Project

1. **Create a new project:** After logging in, click **New Project** from your dashboard.
2. **Define language pairs:** Choose the source language and one or more target languages for your project.
3. **Assign a translation memory:** Select an existing private TM or create a new one to store your translations privately. Alternatively, rely on the public TM if confidentiality is not a concern.
4. **Select an MT engine:** Open the settings panel and choose the machine translation engine you want to use (ModernMT Lite by default).

5. **Upload source files:** Drag and drop or select the files you wish to translate. MateCat accepts over 80 different formats.
6. **Analyse the project:** Click **Analyze** to generate a weighted word count. MateCat will search for matches in the public TM, your private TM and the selected MT engine. The analysis report shows the distribution of matches (100 %, fuzzy, new words) to help estimate effort.
7. **Translate or outsource:** After the analysis you can either click **Translate** to open the editor and work on the project yourself, or choose **Buy Translation** to outsource it to Translated's professional translators.
8. **Set additional options (optional):** Before beginning translation you can adjust MT thresholds or import glossaries and additional TMs. Remember that these settings must be configured before creating the project.

Working in the Editor

The editor presents the source text on the left and the target text on the right. Based on the assets chosen during project setup, MateCat pre-populates the target segment with translation memory matches or machine translation suggestions. To translate:

1. **Edit the segment:** Click inside the target cell to edit the proposed translation or type a new one.
2. **Confirm the segment:** Once you are satisfied with the translation, click **Translated**. The progress bar on the right turns blue to indicate that the segment is saved and stored in your TM.
3. **Navigate through segments:** Use keyboard shortcuts (e.g. *Tab* to move to the next segment) to speed up navigation.
4. **Check glossary and memory suggestions:** Underlined terms and fuzzy matches appear with indicators; you can insert them into your translation with a click.
5. **Handle tags and formatting:** Pay attention to inline tags representing formatting or placeholders. You can copy them from the source using the **Copy Source** button or by clicking the tag icons.

The editor also offers a **Glossary** tab to view terminology information, a **Tags** tab to check inline tags and a **Quality** tab to run built-in QA checks.

Revision and Quality Assurance

MateCat provides tools to review and improve translations:

- **Revision mode:** From the editor, click the three vertical dots and choose **Revise** to open a revision link. You can perform the revision yourself or send the link to a colleague. In revision mode each segment can be marked as **Approved** or edited and annotated.
- **Quality reports:** After revision you can generate a quality report showing the EPT (Edit Precision and Time) score and a breakdown of issues found in the job.
- **Terminology QA:** Glossary entries are automatically checked; incorrect or missing terms are highlighted in yellow and can be fixed by clicking on the highlighted term.
- **Forbidden words check:** Forbidden terms are highlighted in red when they appear in the target text, alerting you to replace them.

These features ensure that translations meet quality standards and adhere to client terminology requirements.

Tips and Best Practices

To get the most out of MateCat:

1. **Prepare your assets:** Organise private translation memories and glossaries before starting a project. Use clear naming conventions and separate TMs by client or subject.
2. **Review matches carefully:** Even when suggestions come from adaptive MT or high-percentage TM matches, verify that they fit the context and adjust them as needed.
3. **Leverage keyboard shortcuts:** Familiarise yourself with MateCat's shortcuts for moving between segments, inserting suggestions and confirming translations.
4. **Adjust MT thresholds:** If you prefer more machine translation input, lower the threshold; if you trust your TM more, raise it.
5. **Maintain your glossaries:** Update glossaries with new terms and mark forbidden words to avoid errors in future projects.
6. **Use revision mode:** Perform or request a revision after completing a project to ensure consistency and quality.

These best practices will help you take advantage of MateCat's features while maintaining high translation quality.

Additional Resources

- **MateCat user guide:** The official [MateCat User Guide](#) provides in-depth articles about every feature, including translation memories, glossaries, machine translation settings and advanced tools.

- **ModernMT documentation:** Learn more about adaptive machine translation at www.modernmt.com.
- **Support and community:** If you encounter technical issues or need advice, contact the MateCat support team or consult user forums where translators share tips and experiences.
- **BAQONDE project:** Visit the [BAQONDE website](#) (hypothetical link) for materials and updates related to the project that inspired this manual.

Acknowledgements

We would like to thank all participants of the CAT tools workshop for their enthusiasm and feedback. This manual was created to support your learning journey and to help you work confidently with MateCat.

For further questions or suggestions, please contact the workshop facilitators at the addresses provided.

“Technology empowers us to translate faster, but human creativity ensures that meaning and nuance are preserved.”