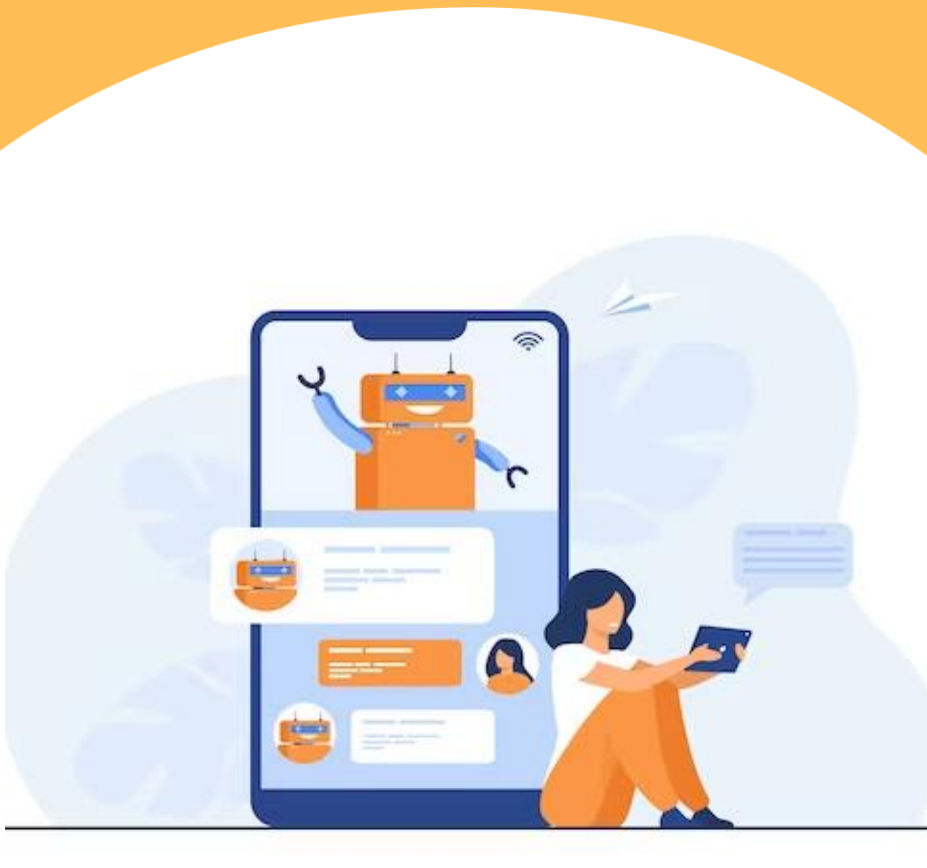




# AI4C PECAT USER GUIDE

**V1.0,  
Oct24**



We are  **Language Technology** 



# INTRODUCTION

AI4C PECAT, a tool designed by Pangeanic to transform machine translation workflows and deliver high-quality outputs.

Here is what PECAT offers to elevate your translation experience:

**Human-in-the-Loop Approach:** PECAT integrates human expertise into the translation process, allowing for the review and refinement of machine-generated translations. This hybrid approach ensures higher accuracy and quality.

**Confidence Score Visibility:** Each translation segment comes with a confidence score, helping you assess the reliability of the machine translation and prioritize segments that require closer attention.

**Advanced TM Management:** Through a user-friendly interface, PECAT provides an advanced Translation Memory (TM) functionality.

You can efficiently access, edit, and utilize existing translation memories, promoting consistency and reducing redundancy and revision time.

**Repetition Auto-Propagation:** PECAT automates the translation of repeated segments throughout your project, saving time and effort by minimizing manual repetition.

**Robust Import/Export Functionalities:** PECAT is integrated with the MINT management system, offering robust import and export capabilities for processing JSON field values. This ensures smooth data exchange and accurate transmission of revised translations.

**Quality Assurance Tools:** PECAT includes spell check to help maintain high standards of translation accuracy and reliability.





**Hotkeys:** Hotkeys are incorporated for quick access to common tasks and functionalities, improving workflow efficiency and user experience.

**Data Utilization for Machine Learning:** PECAT facilitates the creation of parallel data that can be used to train and refine machine translation engines, promoting continuous improvement in translation quality over time.

**Integrated Authentication System:** The tool includes an authentication system integrated with Europeana's platform, ensuring secure and streamlined access for users, and enhancing interoperability with Europeana's infrastructure.

**Context Potential Visualization:** PECAT incorporates context visualization features, providing additional context for translations to aid in producing more accurate and contextually appropriate translations.



1) Access the PECAT platform through its [URL](#), where you will find the homepage (*Figure 1*).

On this first screen, you will find the login page.

Email

jsmith

Password

\*\*\*\*\*

Sign in with Credentials

or

Sign in with Europeana

*Figure 1 Login screen*



# CREATE A NEW PROJECT

1) On the dashboard, you can find a list of projects you created in PECAT in the **Projects** option (*Figure 2*); while using the **Add Project** option, you can create a new project.

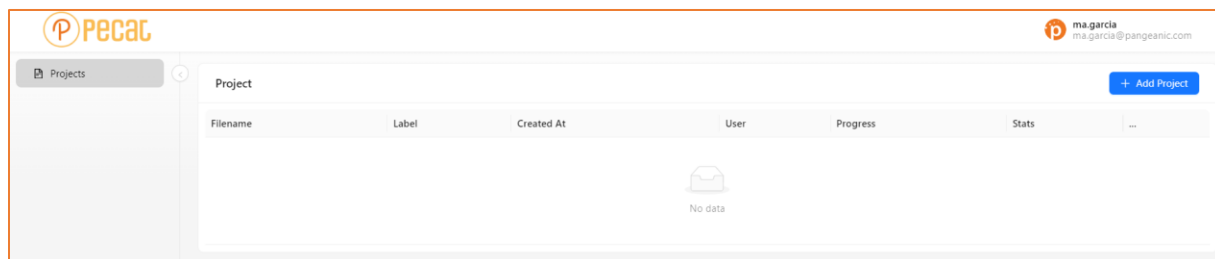


Figure 2. Option

2) A project can be created in two different ways: either by using a URL or by uploading a JSON file, depending on the type of user.

## For CH professionals and Europeana users

Those users who use Mint for loading textual data have the possibility to use a temporary URL, which is created by the user on the Mint platform (**this link lasts 48 hours**).

- The Mint URL will be entered in the **Mint Url** field. (*Figure 3*).
- By clicking on the **OK** button the project is created.

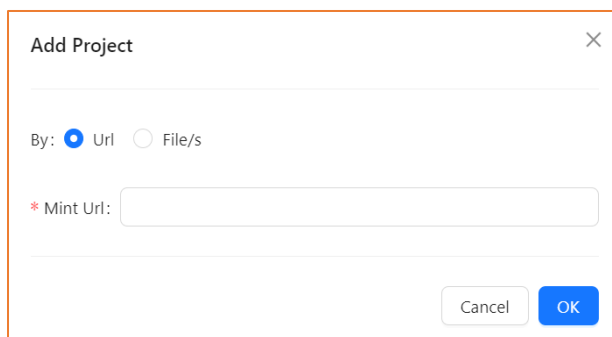


Figure 3. Option Add Projects - URL



# CREATE A NEW PROJECT



## For users who do not use Mint

Users who do not utilize the Mint platform can upload text data in JSON format directly to the PECAT platform.

- In this case, the **File/s** option should be selected on the project creation assistant screen (*Figure 4*).
- A navigation window opens where you can select the JSON file to be loaded.
- Once the file is selected, the project is automatically created and added to the user's dashboard.

Add Project

By: ☐ Url ☒ File/s

File:

*Figure 4. - Option Add Projects - File/s*



# CREATE A NEW PROJECT



## For users who do not use Mint

When loading a file, PECAT allows you to pre-translate the text into English before the project is created. You can optionally select to add MT in the creation process by selecting **Yes** in the **Add TM:** field (*Figure 5*). After selecting **Yes**, select the source and target languages (**the target should be always English since all EuropeanaTranslate's MT engines provide MT into English as a target**) and, then, select the file to be post-edited.

The screenshot shows a dialog box titled "Add Project" with a close button (X) in the top right corner. Inside the dialog, there are several options:

- By:** Two radio buttons, "Url" and "File/s". "File/s" is selected.
- Add MT:** Two radio buttons, "No" and "Yes". "Yes" is selected.
- \* Source language:** A dropdown menu with the text "Select a language" and a downward arrow.
- \* Target language:** A dropdown menu with the text "Select a language" and a downward arrow.
- File:** A button with a download icon and the text "Select File".

Figure 5. - Option Add Projects - File/s





1) As mentioned before, in the user dashboard (*Figure 6*), all the projects created by the user appear. To access a project, click on the file name. In the dashboard, you can also consult various information such as:

- Project creation date
- Progress of the project review
- Number of segments by state type
  - ✓ EDITED: reviewed and corrected segments.
  - ✓ ORIGINAL\_ACCEPTED: reviewed segments where the MT did not need changes.
  - ✓ NOT\_REVIEWED: unreviewed segments.
  - ✓ REJECTED: Reviewed segments that have been rejected as not valid.

Other functionalities offered by the dashboard are:

- File download 
- Project edition 
- Project deletion 

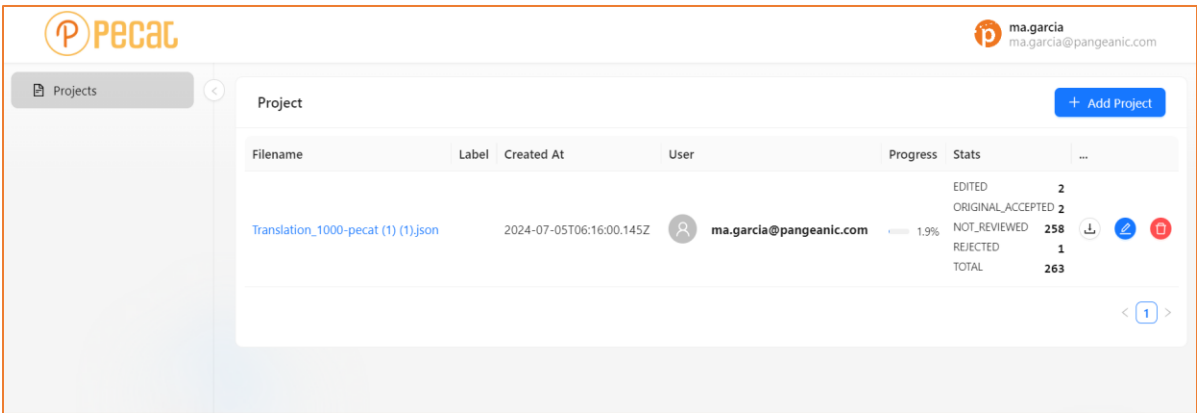


Figure 6. User's dashboard





## EDIT A PROJECT

- 1) The text to be reviewed is presented in the form of a bilingual segment accompanied by metadata that makes it possible to search and filter according to different criteria, such as, for example, the status of the segment in the **Status** column (REJECTED, EDITED, ORIGINAL\_ACCEPTED, or NOT\_REVIEWED), or the confidence coefficient of the MT indicated in the **SP** (Score Percentage) column and can be sorted in ascending or descending order.
- 2) To start editing the machine translation, simply click the target segment.
- 3) Upon completion of the MT editing/revision, the segment is confirmed using the Ctrl+Enter keys or by clicking the **Confirm** button.
- 4) If the segment is not valid (e.g. segment with only spaces, only symbols created by an OCR, etc.), it is possible to reject the segment so that it is not included in the final revised file. This reject action is carried out by clicking on the **Reject** button.

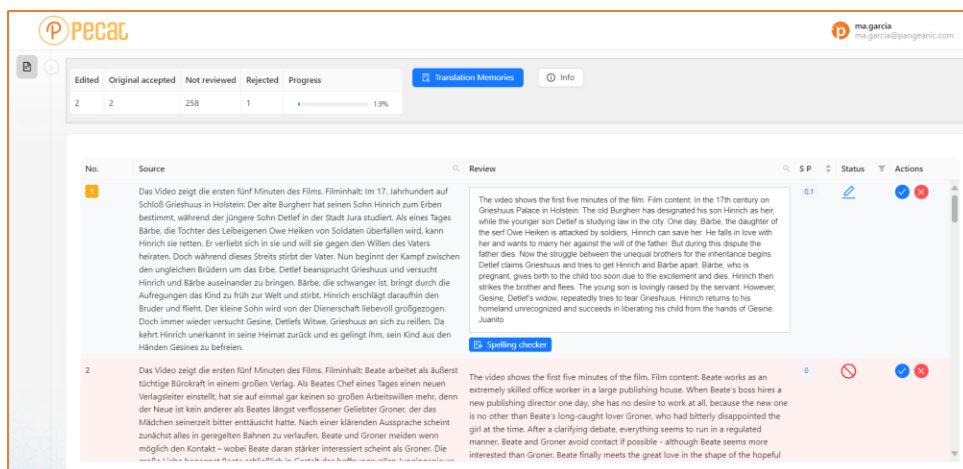


Figure 7. Edition screen





5) On the edit screen, you can check the status of the task review by consulting the progress bar (*Figure 8*).

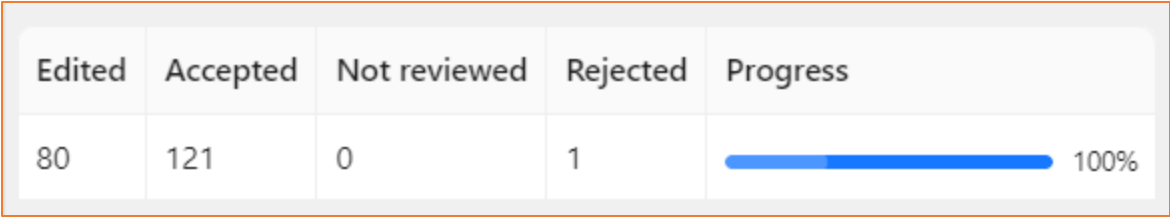



Figure 8. Progress bar

6) As a revision assistance tool, translation memories (TMs) can be created and managed through the  button.

7) On this screen, you can also find a list of shortcuts under the **Info** button. These shortcuts will boost your productivity by allowing you to manage tasks more quickly.

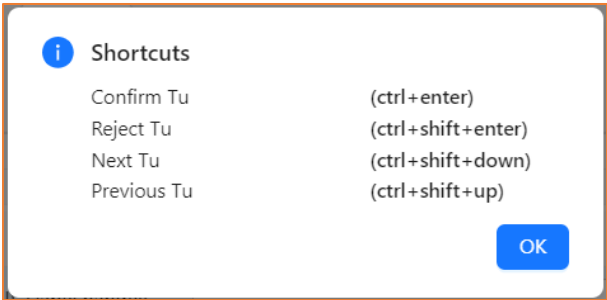


Figure 9. List of shortcuts



# CREATE AND MANAGE TMs

## 3.1. TM creation

- 1) To create a new translation memory (TM), click the **Translation Memories** button.
- 2) In the assistant window, click **Create New TM** (Figure 10) and fill in each field of the displayed form (Figure 11) with the corresponding information, especially those fields that are **mandatory** (\*).

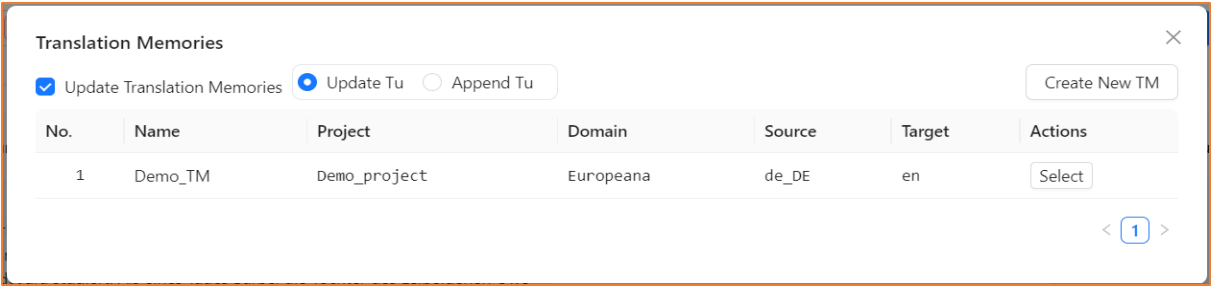
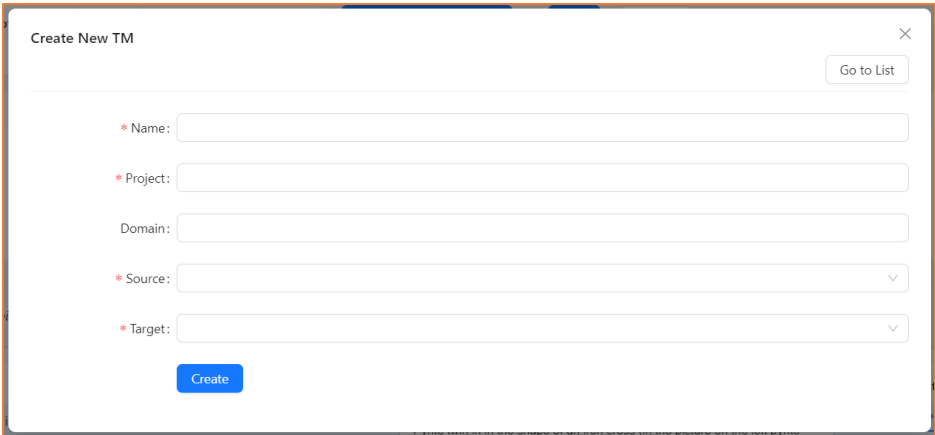


Figure 10. Translation Memories List

- 3) By clicking the **Create** button on the form, the new TM becomes part of the list of translation memories available to this user (Figure 10).



Create New TM

Go to List

\* Name:

\* Project:

Domain:

\* Source:

\* Target:

Create

Figure 11. TM creation screen



## 3.2. TM management

- 1) To select a new translation memory (TM), click on the **TM** button.
- 2) In the assistant window, click **Select** (Figure 10). Once selected, a drop-down window appears, allowing you to consult the TM (Figure 11).

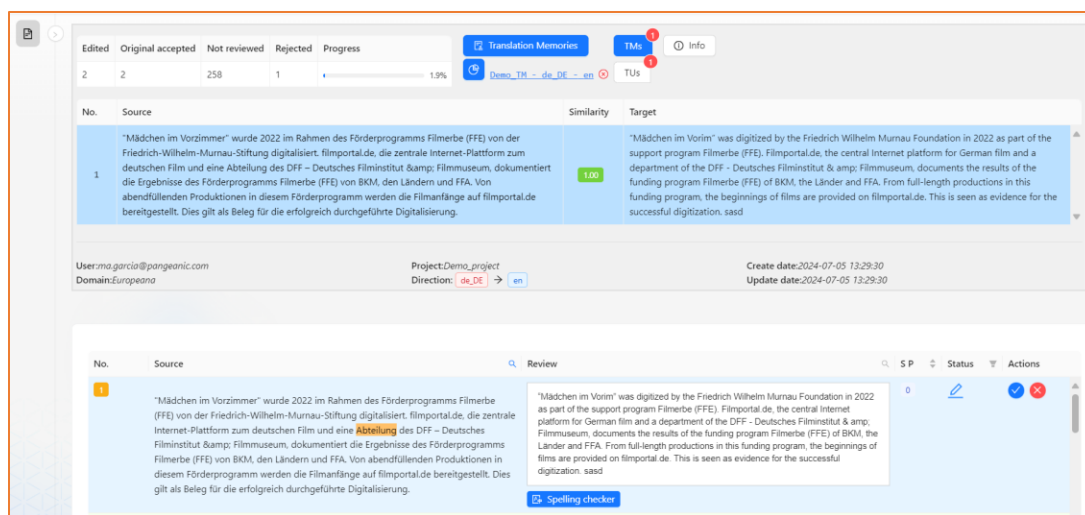


Figure 12. Drop-down window showing the TM matches.

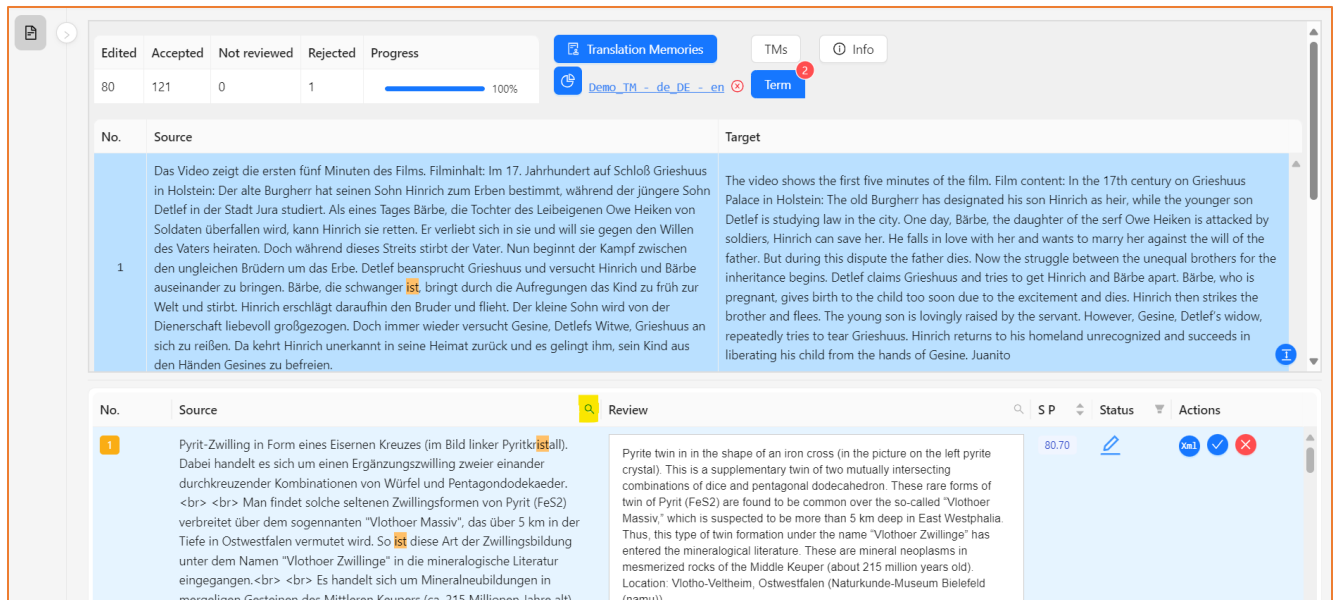
- 3) The translation memory (TM) drop-down window will automatically open whenever the segment being edited has a partial or total match in the TM (Figure 12).



# CREATE AND MANAGE TMs

## 3.3. Terminological consultation in the TM

1) To search for a specific term in the TM, click on the magnifying glass icon (*Figure 14*) and enter the term. Then, click the **Search** button. The TM query window will open automatically, displaying all instances where that term appears in the segments stored in the TM (*Figure 13*).



No.	Source	Target
1	Das Video zeigt die ersten fünf Minuten des Films. Filminhalt: Im 17. Jahrhundert auf Schloß Grieshuus in Holstein: Der alte Burgherr hat seinen Sohn Hinrich zum Erben bestimmt, während der jüngere Sohn Detlef in der Stadt Jura studiert. Als eines Tages Bärbe, die Tochter des Leibeigenen Owe Heiken von Soldaten überfallen wird, kann Hinrich sie retten. Er verliebt sich in sie und will sie gegen den Willen des Vaters heiraten. Doch während dieses Streits stirbt der Vater. Nun beginnt der Kampf zwischen den ungleichen Brüdern um das Erbe. Detlef beansprucht Grieshuus und versucht Hinrich und Bärbe auseinander zu bringen. Bärbe, die schwanger ist, bringt durch die Aufregungen das Kind zu früh zur Welt und stirbt. Hinrich erschlägt daraufhin den Bruder und flieht. Der kleine Sohn wird von der Dienerschaft liebevoll großgezogen. Doch immer wieder versucht Gesine, Detlefs Witwe, Grieshuus an sich zu reißen. Da kehrt Hinrich unerkannt in seine Heimat zurück und es gelingt ihm, sein Kind aus den Händen Gesines zu befreien.	The video shows the first five minutes of the film. Film content: In the 17th century on Grieshuus Palace in Holstein: The old Burgherr has designated his son Hinrich as heir, while the younger son Detlef is studying law in the city. One day, Bärbe, the daughter of the serf Owe Heiken is attacked by soldiers, Hinrich can save her. He falls in love with her and wants to marry her against the will of the father. But during this dispute the father dies. Now the struggle between the unequal brothers for the inheritance begins. Detlef claims Grieshuus and tries to get Hinrich and Bärbe apart. Bärbe, who is pregnant, gives birth to the child too soon due to the excitement and dies. Hinrich then strikes the brother and flees. The young son is lovingly raised by the servant. However, Gesine, Detlef's widow, repeatedly tries to tear Grieshuus. Hinrich returns to his homeland unrecognized and succeeds in liberating his child from the hands of Gesine. Juanito

No.	Source	Review	SP	Status	Actions
1	Pyrit-Zwilling in Form eines Eisernen Kreuzes (im Bild linker Pyritkristall). Dabei handelt es sich um einen Ergänzungszwilling zweier einander durchkreuzender Kombinationen von Würfel und Pentagondodekaeder.     Man findet solche seltenen Zwillingformen von Pyrit (FeS2) verbreitet über dem sogenannten "Vlothoer Massiv", das über 5 km in der Tiefe in Ostwestfalen vermutet wird. So ist diese Art der Zwillingbildung unter dem Namen "Vlothoer Zwillinge" in die mineralogische Literatur eingegangen.    Es handelt sich um Mineralneubildungen in mesozänen Gesteinen des Mittleren Keupers (ca. 215 Millionen Jahre alt).	Pyrite twin in in the shape of an iron cross (in the picture on the left pyrite crystal). This is a supplementary twin of two mutually intersecting combinations of dice and pentagonal dodecahedron. These rare forms of twin of Pyrit (FeS2) are found to be common over the so-called "Vlothoer Massiv" which is suspected to be more than 5 km deep in East Westphalia. Thus, this type of twin formation under the name "Vlothoer Zwillinge" has entered the mineralogical literature. These are mineral neoplasms in mesmerized rocks of the Middle Keuper (about 215 million years old). Location: Vlotho-Veltheim, Ostwestfalen (Naturkunde-Museum Bielefeld (namu)).	80.70		Final ✓ ✗

Figure 13. TM Terminological Search.

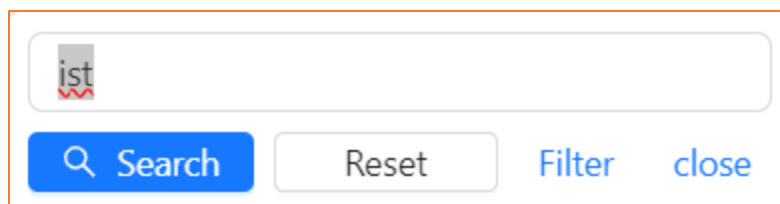


Figure 14. Search assistant (magnifying glass).



# SPELLING REVIEW

1) PECAT allows for an automatic spelling check. Their application is very simple. Once the MT has been reviewed, click on the **Spelling checker** button.

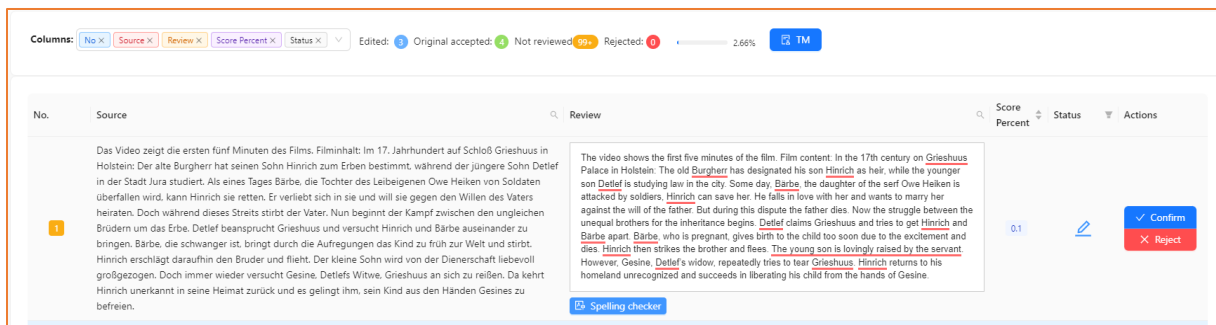


Figure 14. - Spell check.

2) Candidate terms identified as errors will be highlighted with a red line (Figure 14), so that the user can review and make the necessary corrections.

\* By clicking the left mouse button, a correction suggestion will appear, if the system has one.

3) Once the correction is completed, the segment must be confirmed to save the changes.



1) PECAT includes a context visualization feature that allows users to view the original EDM file (MINT's preferred format for metadata storage). This feature is enabled by clicking on the **Xml** button (*Figure 15*) on the right side of the segment.

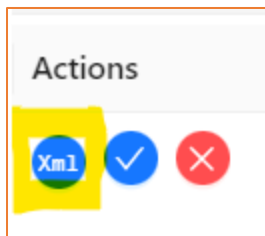


Figure 15. - Xml button.

2) The original EDM is displayed every time the user clicks on the **Xml** button (*Figure 16*). However, the EDM will be available for visualization during 90 days. After that period, the link access will expire.

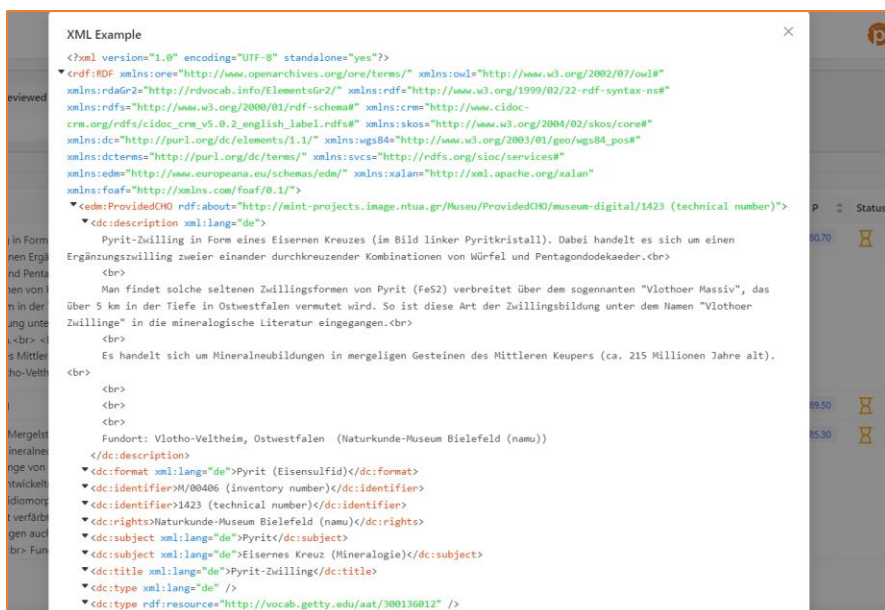


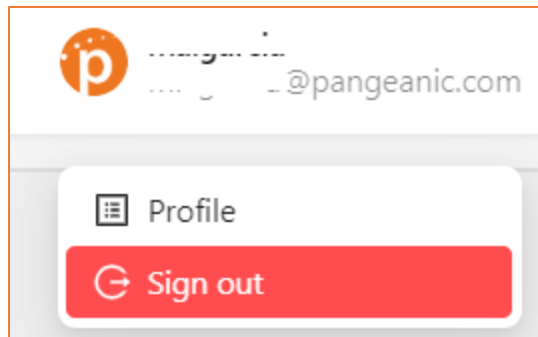
Figure 16. - EDM visualization.



## LOG OUT



1) To quit the platform, click on the user's name that appears in the upper right corner. By doing so, the **Sign out** option (*Figure 17*) will appear, allowing you to exit the system.



*Figure 17. - Sign out option.*





# FAQ

**I have confirmed a segment and it has not been saved to the translation memory.**

- Refresh your browser page. The changes may not have taken effect due to the connection.

**What file formats does PECAT accept?**

- "pdf", "txt", "docx", "docm", "dotx", "dotm", "pptx", "pptm", "potx", "potm", "ppsx", "ppsm", "xlsx", "xlsm", "xltx", "xltm", "vsdx", "vsdm", "ods", "ots", "odt", "ott", "odp", "otp", "odg", "otg", "po", "idml", "json".

Please contact our assistance service at: [support@pangeanic.com](mailto:support@pangeanic.com) to send queries or report technical issues related to PECAT.

