



How To Use Transkribus in 10 Steps

🔗 **Basic overview** 📄 Transkribus Expert Client

Last update 6 months ago

« [How-To's overview \(https://readcoop.eu/transkribus/resources/how-to-guides/\)](https://readcoop.eu/transkribus/resources/how-to-guides/)

► About Transkribus ()

Table of Contents	^
How To Use Transkribus in 10 Steps	
1 – Introduction	
2 – To use Transkribus – register at the website	
3 – Download Transkribus from the website	
4 – Open Transkribus	
5 – Upload your documents	
6 – Segment your documents into lines	
7 – Start your transcription	
8 – Save and export your transcription	
9 – Use Handwritten Text Recognition (HTR) on your documents	
10 – Searching	
Credits	

This document is a basic introduction to Transkribus. It provides a simple standard workflow for working with the platform. If you need more detailed instructions on the functions of Transkribus, please have a look at our other How to Guides (<https://readcoop.eu/transkribus/knowledge-base/how-to-guides/>).

1 – Introduction

Transkribus can be used for several purposes. The most important are:

- Transcribe documents for a scholarly edition
- Create training data to feed the **Handwritten Text Recognition (HTR+)** system so it can learn to decipher your historical documents.
- Run HTR+ on your documents and receive **automatically generated transcripts**.
- Search for distinct words in your document collections.
- The platform lives from the community. The more data uploaded to Transkribus, the more efficient the program and especially the Handwritten Text Recognition will get

Transkribus is a research infrastructure which was established as part of the H2020 Project **READ (Recognition and Enrichment of Archival Documents)**.

Take some time to explore Transkribus and become familiar with how it works. To make it easier, we have created several How to Guides, which give instructions on the different functions of the platform. You can find them within our Resource Base.

(<https://readcoop.eu/transkribus/knowledge-base/how-to-guides/>)





Account

(https://readcoop.eu)

2 – To use Transkribus – register at the website



- Go to our Transkribus page (<https://readcoop.eu/transkribus/>)
- Read and accept our Terms & Conditions (<https://readcoop.eu/terms-and-conditions/>) and our Privacy Policy (<https://readcoop.eu/privacy-policy/>)
- All documents uploaded to Transkribus are “private”, which means that no one except you has access to them.
- The Transkribus team fully supports all EU directives on data protection and privacy. We will respect your privacy and only use the data you give us to improve our services and support future research in humanities and computer science!

3 – Download Transkribus from the website

- Go to the Transkribus download page (<https://readcoop.eu/transkribus/download/>)
- Transkribus runs on Windows, MacOS and Linux. If you need help installing the platform, you can have a look here (<https://readcoop.eu/transkribus/wiki/download-and-installation/>).
- If you use MacOS an error message may appear when you try to open Transkribus for the first time. To remedy this:
 - right click the Track Pad to open the Context Menu and add a security exception for Transkribus.
- Once you have downloaded Transkribus, make sure you unzip the file. The program cannot be started from the zipped file.

4 – Open Transkribus

- Start the tool and use the “Login” button in the “Server” tab.

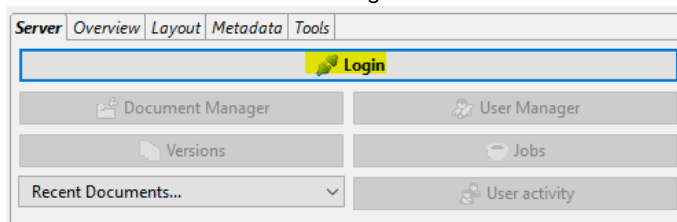


Figure 1. Login

- You will have access to your private collection named after your email address. This collection includes some test documents that you can experiment with.
- You can find it by clicking the “Collections” button in the “Server” tab.

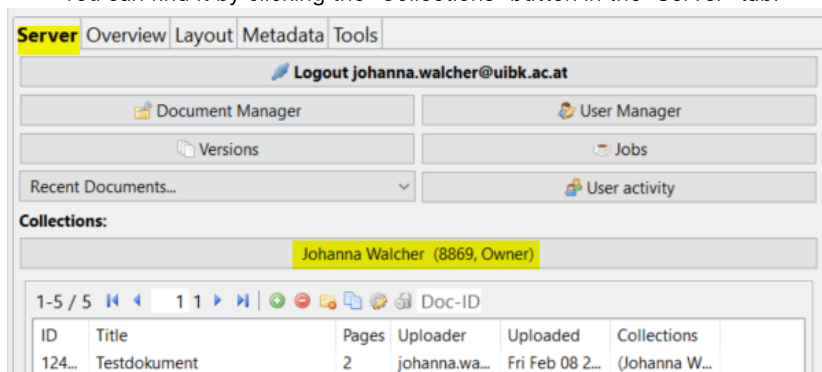


Figure 2. Test documents in your collection



Account ▾

5 – Upload your documents



- (<https://readcoop.eu>)
- Transkribus allows you to work with your own documents, either locally or by uploading them to the server.
 - Automated processes can only be performed if the documents are uploaded to the Transkribus platform. The platform can process PDF, JPEG, PNG and TIFF files. JP2 files are not supported unfortunately.
 - You can upload documents which you have scanned yourself. You can also use our **DocScan app** for Android smartphones to take images and upload them directly to Transkribus. For more information: <https://scantent.cvl.tuwien.ac.at/en> (<https://scantent.cvl.tuwien.ac.at/en>)
 - You may also download documents from the Internet and upload them to Transkribus. Many libraries and archives follow Open Access policies and are therefore encouraging further usage of their collection – you can ask archives and libraries directly if you can upload images of their documents to Transkribus!
 - Click the “Import document(s)” button to transfer the images from your computer to the platform. **Note:** the images need to reside in a separate folder on your computer before you upload them to Transkribus!



Figure 3. Upload your documents to Transkribus

- You can add your documents to one of your existing collections or create a new one by clicking the “Add to collection” button at the bottom of the “Document ingest/upload” box and then clicking “Create”.

Document ingest / upload

☐ Upload via private FTP (also PDF files)
 ☒ Upload single document

☐ Upload via URL of DFG Viewer METS
 ☐ Upload via URL of IIIF manifest

☐ Extract and upload images from pdf

Single document upload

Local folder:

Title on server:

Add to collection: How-To Guides (148300, Owner)

Figure 4. Add documents to one of the existing collections or create a new one

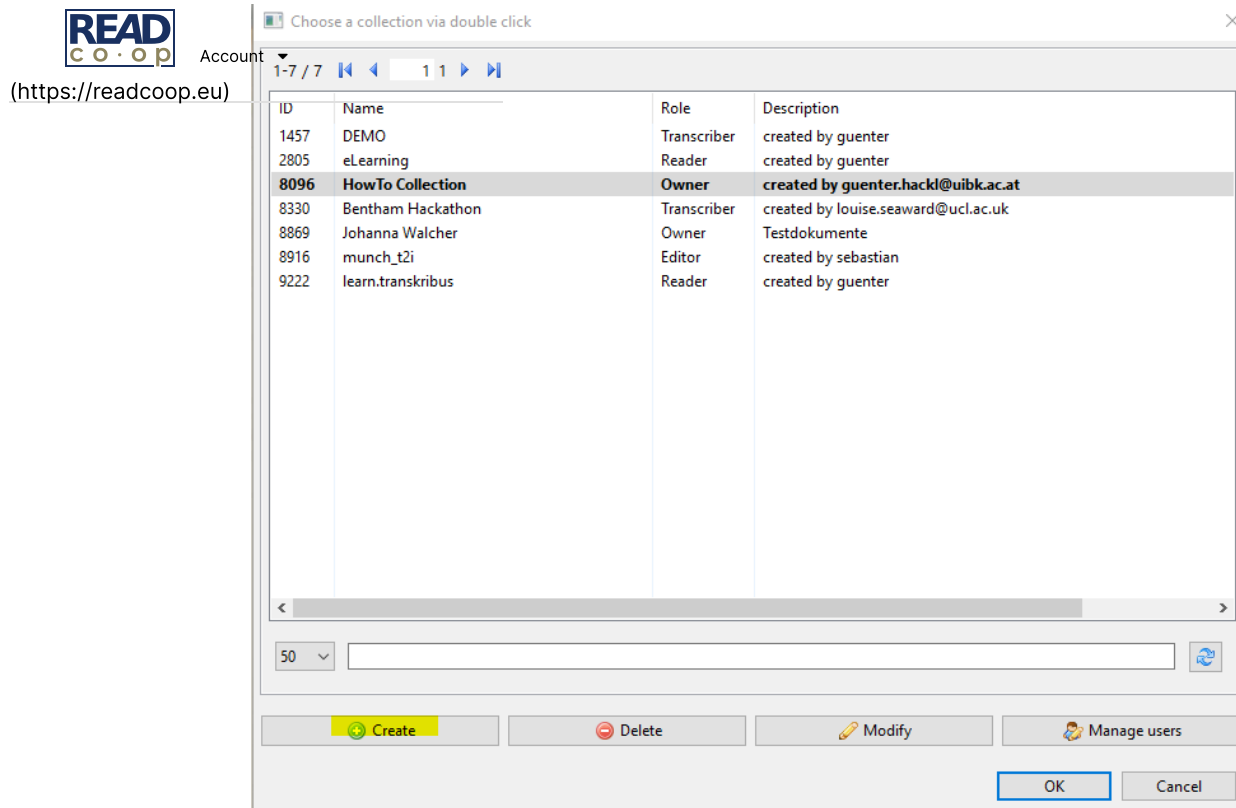


Figure 5. Create your own collection

- To access your documents, click on the “Collections” button in the “Server” tab and choose your collection. Then double-click on the documents in the box at the bottom of the “Server” tab to open them.

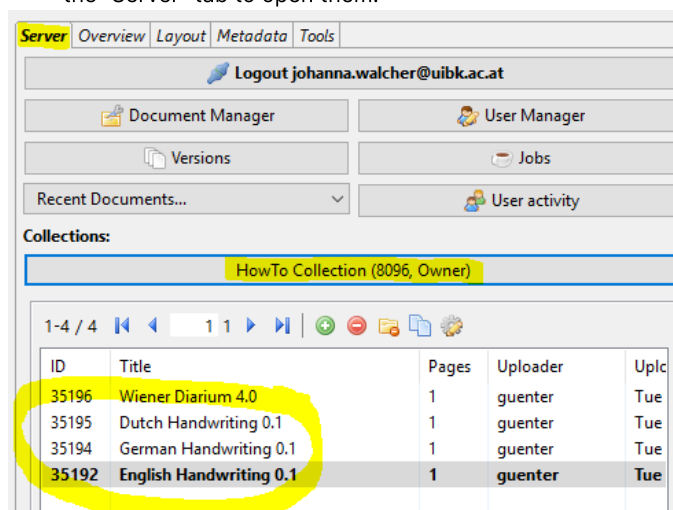


Figure 6. Open the documents in your collection

- All documents uploaded to Transkribus are private by default.** You can give other users authorisation to view your documents if you wish. Use the “User Manager” button in the “Server” tab to add users to your collection. You can only share collections with users who have a Transkribus account.

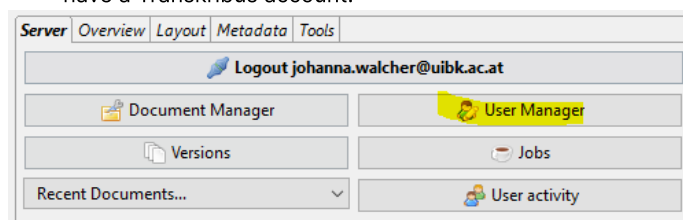


Figure 7. “User Manager” button for managing access to your collection



Account ▾

6 – Segment your documents into lines



- (<https://readcoop.eu>)
- To feed the HTR engine with training data your documents need to be segmented into lines. This can be done automatically in Transkribus.
 - Open the “Tools” tab and go to the Layout Analysis section.
 - You can choose to segment the current page or a batch of pages.
 - Make sure “Find Text Regions” is selected and press “Run”.
 - The lines and text regions in your document will be detected automatically.
 - If no or few lines are detected, your images have a low resolution, or too many text regions have been created, read here ([▼ **Layout Analysis**

Method: Transkribus LA Configure...

☒ Current page

☐ Pages \(18\): 1-18

Filter by: NEW,IN_PROGRESS

▼ **Document Selection**

Test Baselines Detection_duplicated \(1281361\)

☐ Current collection Choose docs...

☒ Find Text Regions ☒ Find Lines in Text Regions

Model: Preset

Run](https://readcoop.eu/transkribus/howto/how-to-transcribe-documents-with-transkribus-introduction/#:~:text=and%20transcription%20tasks-,Automatically%20detect%20text%20regions%2C%20lines%20%20automatically%20run) how to configure the Layout Analysis settings to get better results.

</div>
<div data-bbox=)

Figure 8. Layout Analysis

7 – Start your transcription

- Once the baselines are visible on your image, you can write text into the Text Editor field.
- For each baseline, there will be a **corresponding line in the Text Editor**. Transcribe the text line by line exactly as it appears in the image.

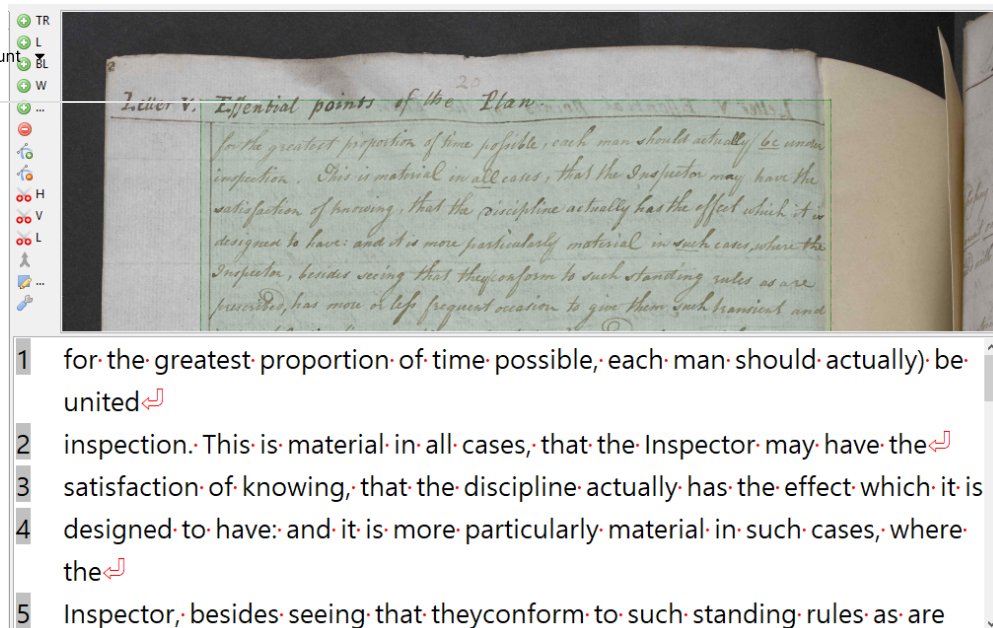


Figure 9. Transcription

- Special characters can be found in the “Virtual Keyboards” button in the Formatting toolbar below the Text Editor.

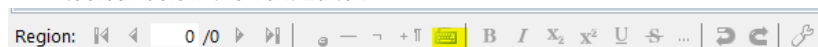


Figure 10. “Virtual Keyboards” button

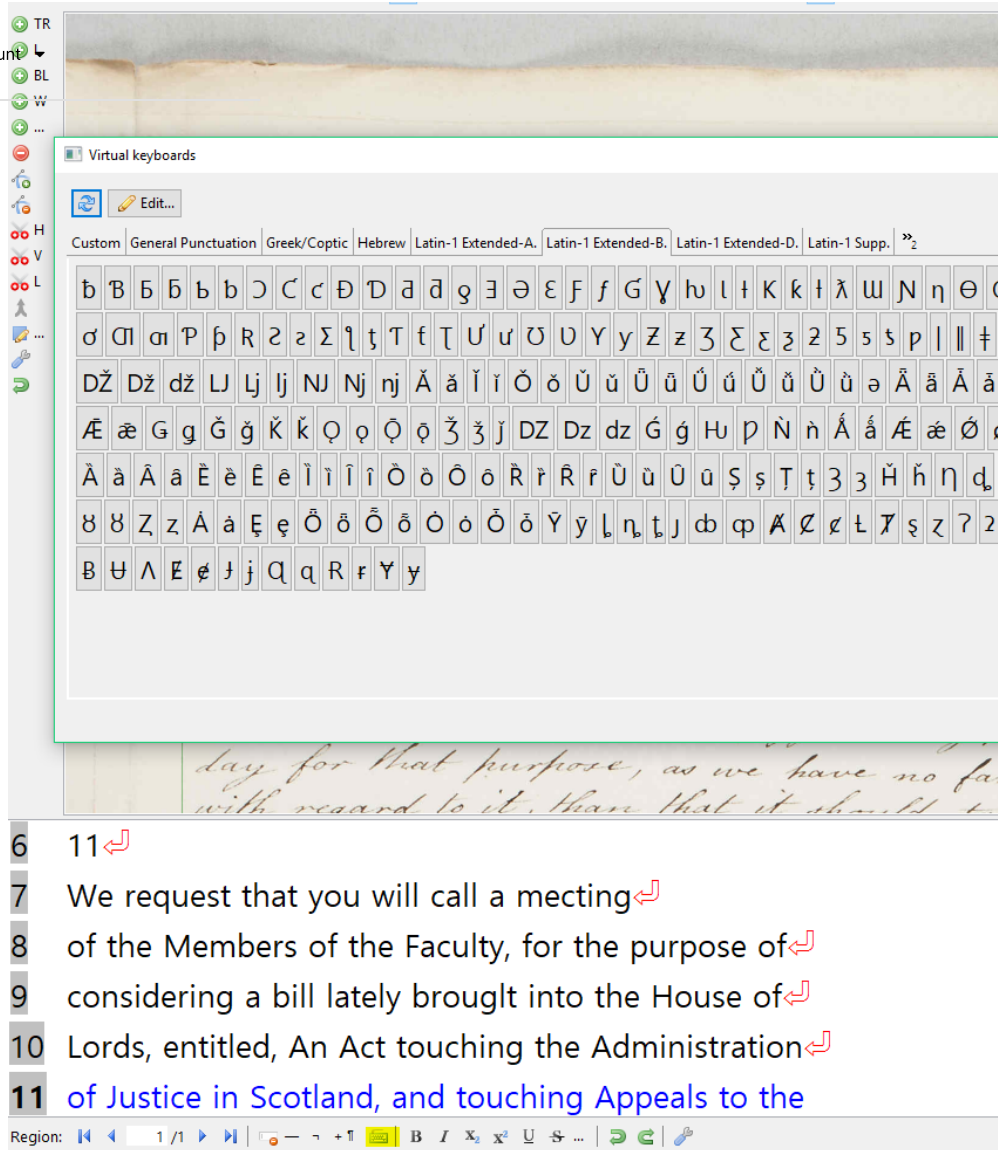


Figure 11. Virtual keyboards

- If you are working in a team, you might find it easier to transcribe in **the browser version of Transkribus, Transkribus Lite** (<https://lite.transkribus.eu/>).

8 – Save and export your transcription

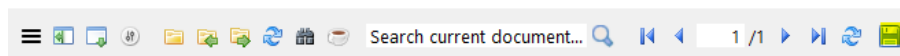


Figure 12. Saving the changes in your document

- Press the “Save” button in the Main Menu to save the document in Transkribus.
- If you click on the “Versions” button in the “Server” tab, you will see that a **new version** has been created. This means that you can always access previous versions of a document should you need to.

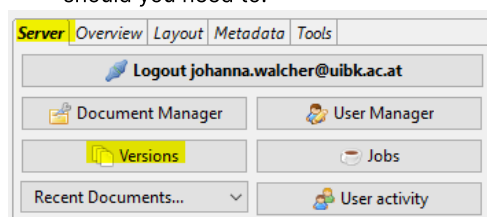


Figure 13. Click the “Versions” button to access previous versions of your document

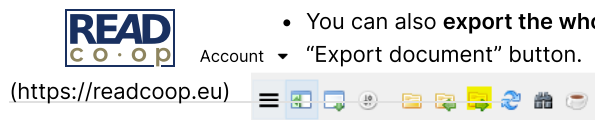


Figure 14. "Export document" button

9 – Use Handwritten Text Recognition (HTR) on your documents

- It is simple to have your documents recognised by the computer. You can use one of the public models (<https://readcoop.eu/transkribus/public-models/>) made available by the Transkribus community, or you can start training your own model: to do so, you need around 5,000 transcribed words of printed text or 15,000 words of handwritten text.
- Once an **HTR model** has been trained for your documents, it can be applied via the "Run" button in the "Text Recognition" section in the "Tools" tab. You can select one or more pages of your documents and start recognition.

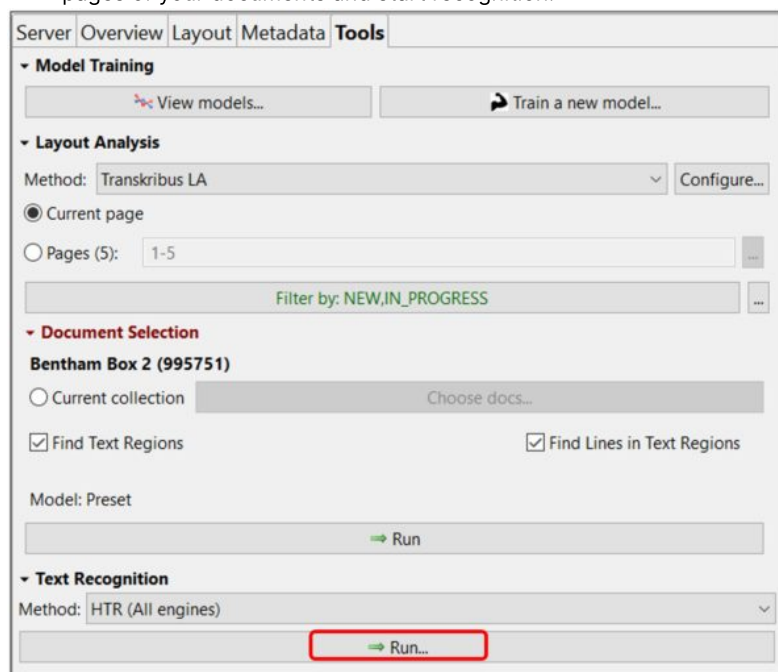


Figure 15. Run Handwritten Text Recognition

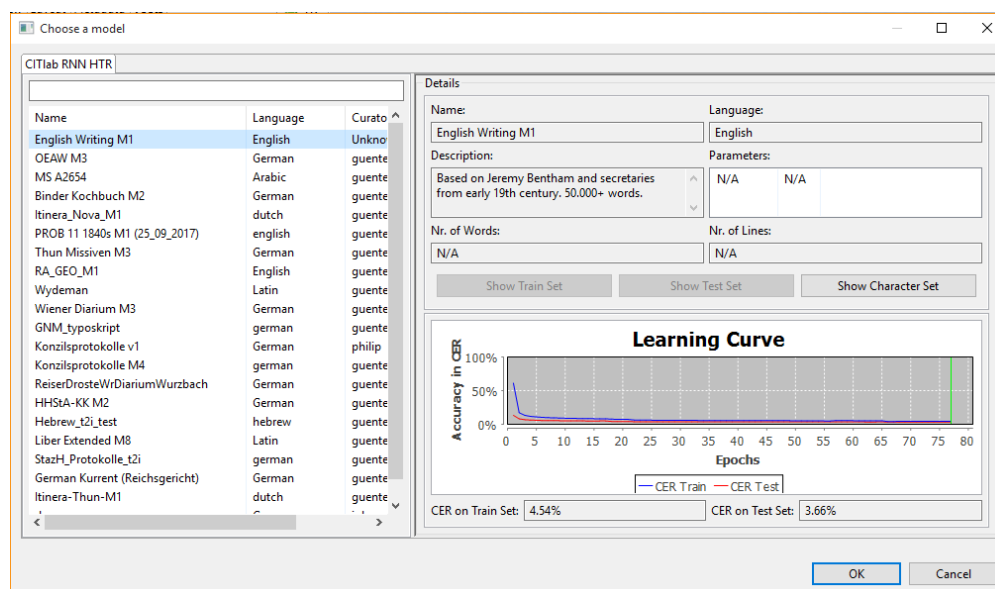


Figure 16. Model overview and learning curve



Account

(https://readcoop.eu)

- If you click “Run” and then “Select HTR model”, you can choose the model for the recognition and get more information about it.
- On the left side of the window, you can see an overview of the available models.
- On the top right side of the window, the details of the model are shown.
- The graph on the bottom right signifies the accuracy of your model with the Character Error Rate (CER), i.e. the percentage of characters that have been transcribed incorrectly by HTR. The blue line represents the progress of the training. The red line represents the progress of evaluations on the Test Set of data which was set aside during the training process.
- After the HTR has finished, the results will appear directly on a new version of your document within Transkribus. It is possible to **evaluate the accuracy** of the automatic transcription using the “Compute Accuracy”-function in the “Tools” tab.



10 – Searching

- Once you have run the HTR model on your documents, you can search for specific terms.
- Open the Search function with the binoculars button shown in Figure 18.



Figure 18. Open the “Search for...” window

- In the window which opens up, choose the “Fulltext (Sorl)” tab.
- Simply type the word you would like to search and press the “Search” button.
- The results will appear below: by hovering with the cursor on a result, you will see the word preview; when you double-click on a result, the page containing it will open in the background.
- You can also flag the “Fuzzy Search” option: Fuzzy Search is a search technique that allows finding similar words in addition to exact matches for the search term. This can come in handy when alternative spellings are plausible.

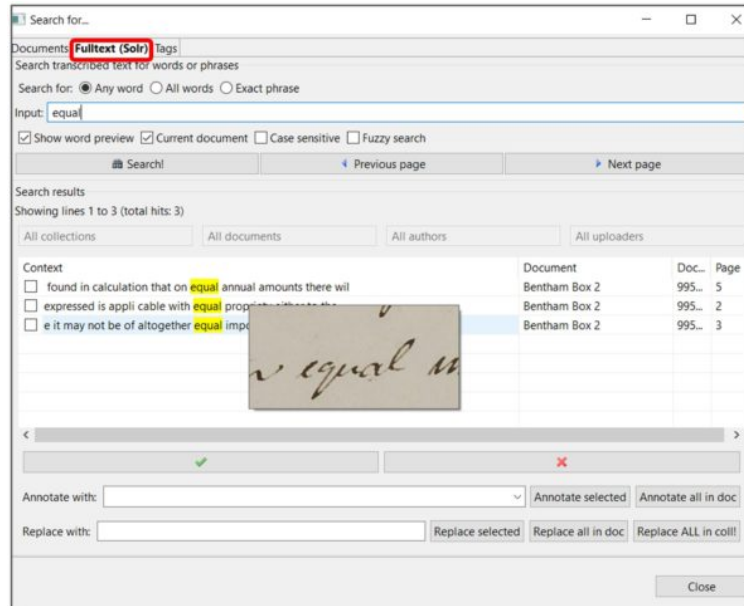


Figure 19. Fulltext Search

Credits

We would like to thank the many users who have contributed their feedback to help improve the Transkribus software.

Transkribus and the technology behind it are made available via the following projects and sites:



Account

- <https://readcoop.eu/> (<https://readcoop.eu/>)
- <https://github.com/transkribus/> (<https://github.com/transkribus/>)



(<https://readcoop.eu/>) **Contact**

- The Transkribus Team: info@readcoop.eu (<mailto:info@readcoop.eu>)



The Transkribus Platform is provided by the European Cooperative **READ-COOP SCE**.

Until June 2019 Transkribus was financed as part of the Horizon 2020 READ-project under grant agreement No. 674943.

The COOP

About us (<https://readcoop.eu/about/>)

Join us! (<https://readcoop.eu/join/>)

Our Members (<https://readcoop.eu/members/>)

Success Stories (<https://readcoop.eu/success-stories/>)

Work with us (<https://readcoop.eu/work-with-us/>)

Products & Services

Transkribus (<https://readcoop.eu/transkribus/>)

Transkribus lite (<https://transkribus.eu/lite>)

Read&search (<https://readcoop.eu/readsearch/>)

ScanTent (<https://readcoop.eu/scantent/>)

Useful information

News (<https://readcoop.eu/news/>)

Download Transkribus (<https://readcoop.eu/transkribus/download/>)

Public Models (<https://readcoop.eu/transkribus/public-models/>)

Payment and shipping (<https://readcoop.eu/payment-and-shipping/>)

Helpful resources

Resource center (<https://readcoop.eu/transkribus/resources/>)

How-to Guides (<https://readcoop.eu/transkribus/resources/how-to-guides/>)

Getting started with Transkribus (<https://readcoop.eu/transkribus/start/>)

FAQs (<https://readcoop.eu/transkribus/questions/>)

Videos (<https://readcoop.eu/transkribus/resources/videos/>)

link 'w
ps:/ edi e.c ww.
/twCommunity fac
tter n.c cha ebo
co om/ nne
m/T inco l/U ok.
ran mp C- co
skri any txV m/t
bus /rea gM ran
) d- 31r skri
coo DT bus
p- GIB)
sce nH-
) zpP
iA)

Copyright © 2023 READ-COOP SCE

- [Terms & conditions \(https://readcoop.eu/terms-and-conditions/\)](https://readcoop.eu/terms-and-conditions/)
- [Privacy Policy \(https://readcoop.eu/privacy-policy/\)](https://readcoop.eu/privacy-policy/)
- [Contact \(https://readcoop.eu/contact/\)](https://readcoop.eu/contact/)
- [Imprint \(https://readcoop.eu/imprint/\)](https://readcoop.eu/imprint/)

EN