

<https://readcoop.eu>[Transkribus \(/transkribus/\) >](#)[How-To Guides \(https://readcoop.eu/transkribus/resources/how-to-guides/\) >](https://readcoop.eu/transkribus/resources/how-to-guides/)

How to compute accuracy of HTR models

How to compute accuracy of HTR models

[🔍 Transkribus Tools](#) [📋 Transkribus Expert Client](#)*Last update 1 year 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 compute accuracy of HTR models

Introduction

[Compute accuracy](#)

Options to check the results of automated transcriptions

Compare text versions

Compare



Advanced compare

Compare Samples

Introduction

This guide will show you how to compute the accuracy of different models and compare the accuracy of the automatic transcription in more detail as well as how to compare the accuracy of a recognition model run on a sample set of your specific material.

Compute accuracy

You can measure the accuracy of your model on specific pages of your Training and Validation Sets with the “Compute Accuracy” feature in the “Tools” tab. To do so, you first need to generate an HTR transcript.  

(<https://readcoop.eu>)

As “**Reference**”, choose a page version that has been transcribed correctly (Ground Truth: manual transcription as close to the original text as possible). To get the most significant value it would be best to use pages from a sample set which have not been used in the training and therefore are new to the model. Using pages from the Validation Set is also an option but not as ideal as the one just mentioned. Using pages from the Training Set is not a good idea because this will output CER-values that are lower than they actually are.

As “**Hypothesis**”, choose the version that was automatically generated with an HTR-model and on which you would like to test how good the result is.

You can change the versions to be compared by clicking on the grey button next to “Reference” and “Hypothesis”. Double-click to choose the desired version of the document in the appearing window. The versions that can be selected for “Reference” and “Hypothesis” are different versions of your document, which have been created after running a new job or saving transcriptions.



Account

<https://readcoop.eu>

Server Overview Layout Metadata **Tools**

Layout Analysis

Method: CITIab Advanced Configure...

☐ Current page ☒ Pages (6): 1-6 ...

▼ **Document Selection (for admins)**

Copy of HTR Bentham TEST (3743)

☐ Current collection Choose docs...

☒ Find Text Regions Only use unsegmented pages

☒ Find Lines in Text Regions

→ Run

▼ **Text Recognition**

Method: HTR (CITIab HTR+ & PyLaia) ▼

Models... Train...

→ Run...

▼ **Compute Accuracy...**

Reference: 09.10.16 09:44:07 - guenter - Ground Truth Use current

(Correct Text)

Hypothesis: 03.04.17 22:32:59 - tobias.hodel@uzh.ch - In Progress - CITIab HTF

(HTR Text)

Compare Text Versions...

Compare...

Compare Samples...

Figure 1 "Compute Accuracy" within the tab "Tools"

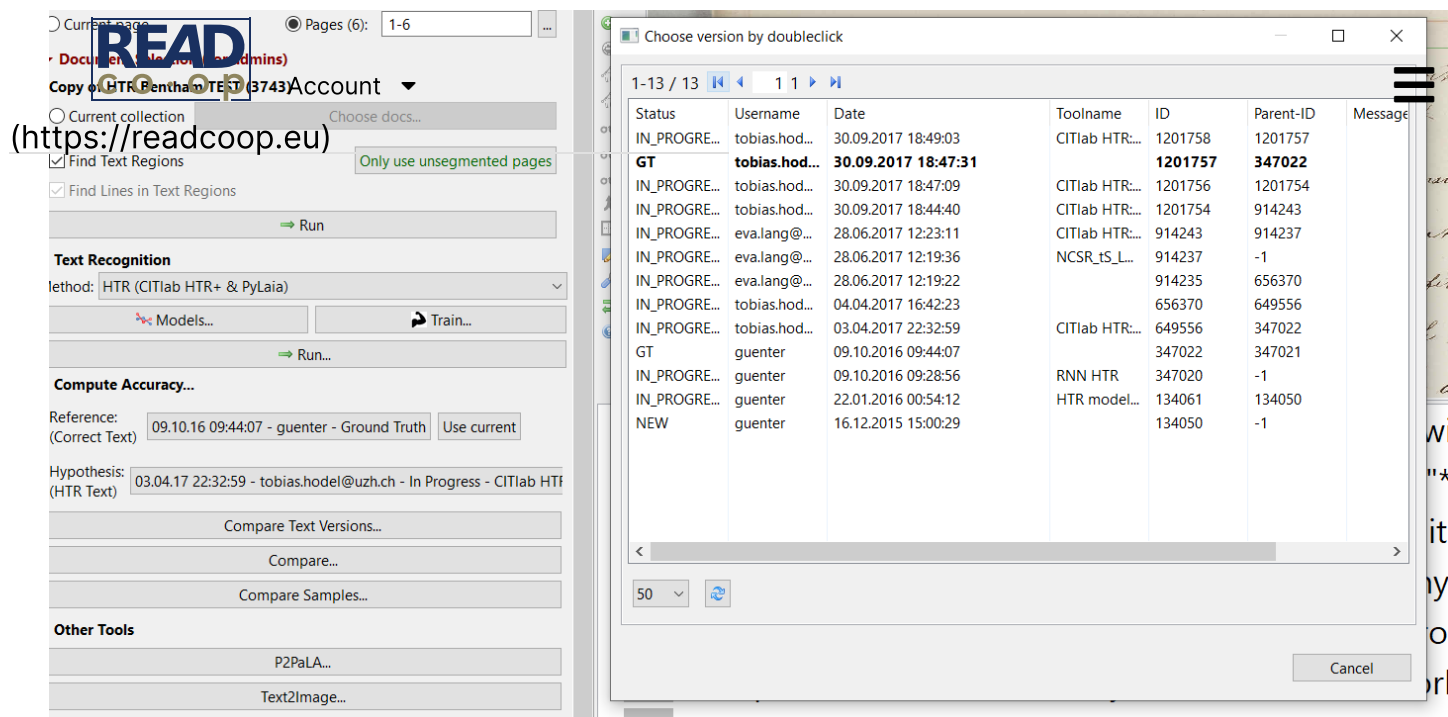



Figure 2 Choosing the right version by double-clicking

Options to check the results of automated transcriptions

Compare text versions

If you click on "Compare Text Versions" you will get a visual representation of what the HTR model transcribed correctly and incorrectly.

READ Comparison Account 

(<https://readcoop.eu>)

Show line numbers

"more than two ~~turnheys~~ turnkeys from morning to night, without
 "allowing them leisure for any other ~~employment:~~ employment."*

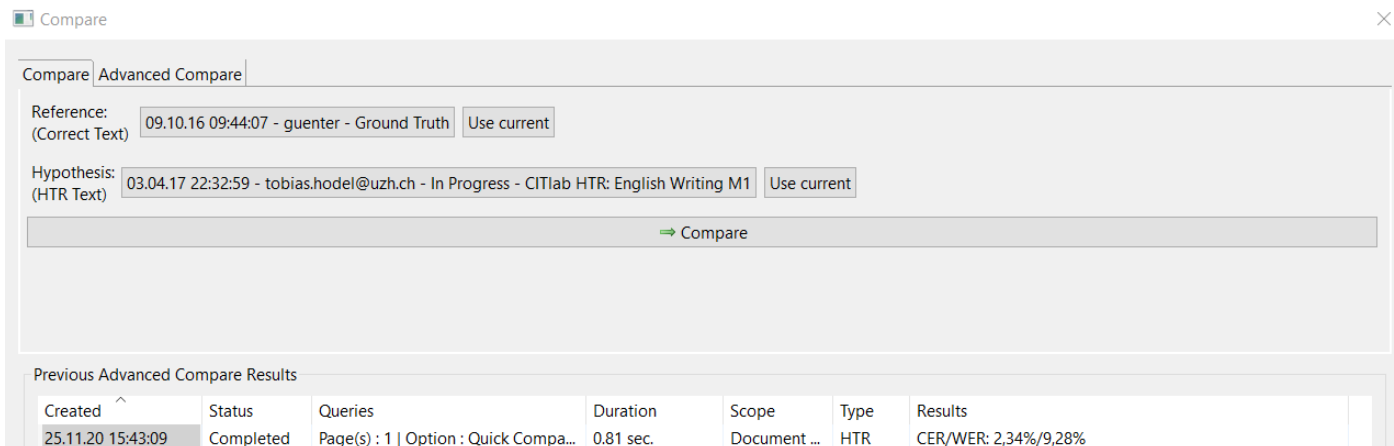
No No. 8. Page 26. "If the undertaking were not unsuitable
 "to my character and habits of ~~life~~ (says life) (says the worthy Magis-
 5 ~~trate~~ -trate)" I would not hesitate to contract for the profits of
 the prisoners "the prisoners' labour, (if they were allowed to work in
 "their respective calling, and in such other labour as is
 "proper for prisoners,) and to pay the County £400 per
 10 ~~annumr.~~ "annum."
 Waving, Waiving, though but for the moment, an ~~observations~~ observation
 15 suggested by the hypothetical offer, expressed by Not, No, I
 come now, my Lord, to the second part of my text. Good
 or not good as it might be for the Counties to be forced
 to ~~corver~~ cover ~~cover~~ themselves with these improved prisons,
 so it is, they will not all of them at least suffer ~~them~~ them-
 -selves to be thus forced. Submitting to the ~~Cox~~ tax ~~do~~ di-
 -rectly -recty and actually imposed upon them, first by his
 Grace, and now by Your ~~Lordship~~ Lordships, for the maintenance
 of prisoners whom Parliament has ordered elsewhere ~~sub~~ - sub-
 -mitting -mitting to this tax, perhaps for what of want of having
 so distinct a view of the noble hands to which they
 are indebted for it as this disclosure will now ~~helps~~ help
 20 them to, they ~~denier~~ denmur beforehand to the so much
 heavier as well as less directly imposed tax, also declared
 to be in store for them _ the tax of loading them, a la
 Vanbrugh, with Blackburn architecture. The tax for
 maintenance
 * A specimen of the ~~expense~~ expence of the official establishments upon
 the plan of the improved prisons especially when on the ~~gully~~ gally

Figure 3 Compare Text Versions

Please note that even if only one character is wrong, the whole word is marked in red. In green, the word is shown as it is written in the Ground Truth transcription. In the passages without colour the recognised text is identical with the Ground Truth.

(<https://readcoop.eu>) is the quickest version. To access it, click on “Compare...”.

Firstly, please make sure the right versions have been selected in the upper section of the appearing window. Then hit the “Compare” button. The result will be shown in the lower section of the window after a few seconds.



The screenshot shows the 'Compare' window with two tabs: 'Compare' and 'Advanced Compare'. The 'Advanced Compare' tab is active. It contains two input fields: 'Reference: (Correct Text)' with the value '09.10.16 09:44:07 - guenter - Ground Truth' and a 'Use current' button; and 'Hypothesis: (HTR Text)' with the value '03.04.17 22:32:59 - tobias.hodel@uzh.ch - In Progress - CITlab HTR: English Writing M1' and a 'Use current' button. Below these is a large 'Compare' button. At the bottom, there is a section titled 'Previous Advanced Compare Results' containing a table with the following data:

Created	Status	Queries	Duration	Scope	Type	Results
25.11.20 15:43:09	Completed	Page(s) : 1 Option : Quick Compa...	0.81 sec.	Document ...	HTR	CER/WER: 2,34%/9,28%

Figure 4 Results

The values are calculated for the page you have currently loaded in the background. In the example image, we have a CER of 2.34% on that page, which means that 97.66% of the characters in the automated transcript are correct.

By double clicking the date and time within the “Created” column of the simple comparison tab, you will automatically arrive at the “Advanced Statistics” window. Here you will get more detailed indications and values and the results can be exported into an Excel file.

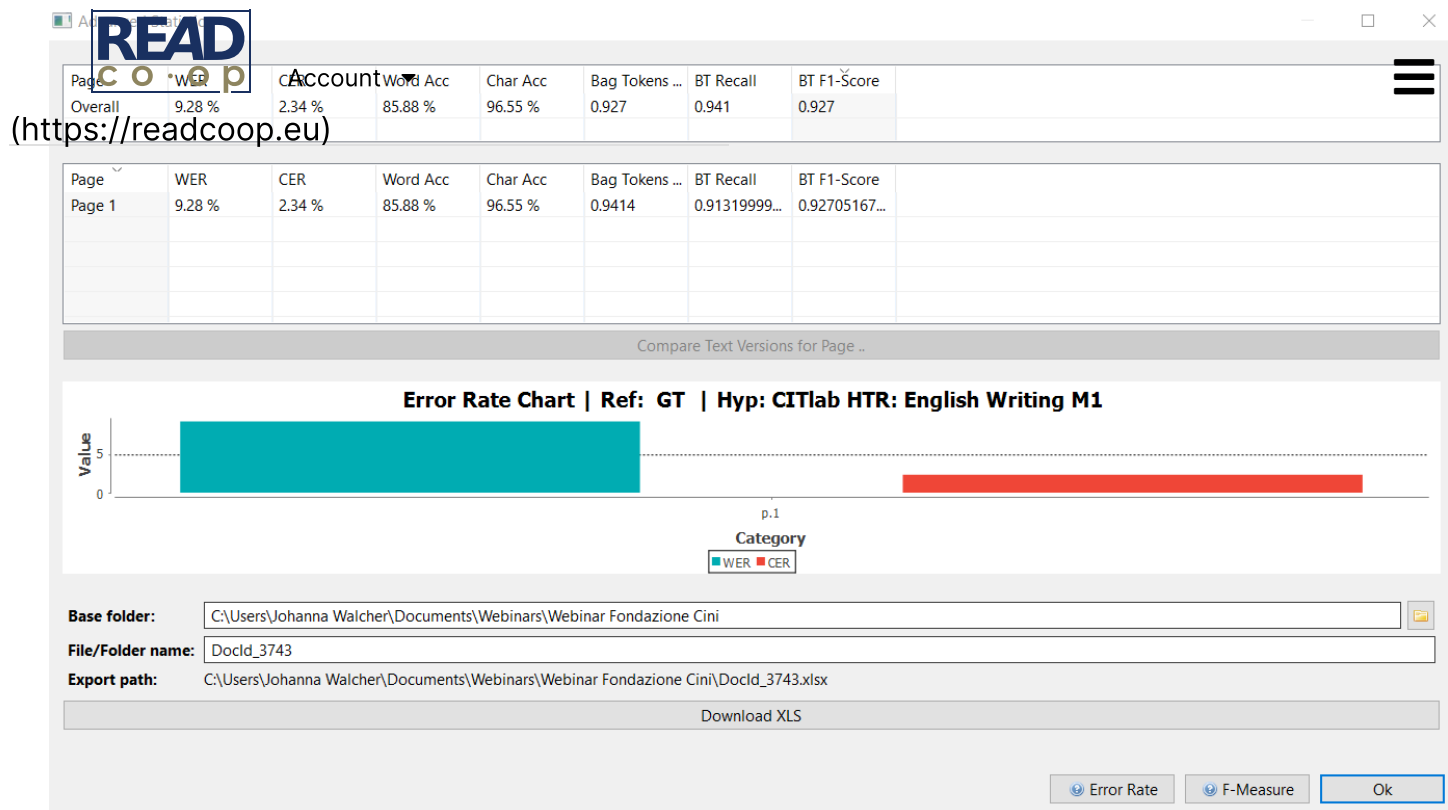


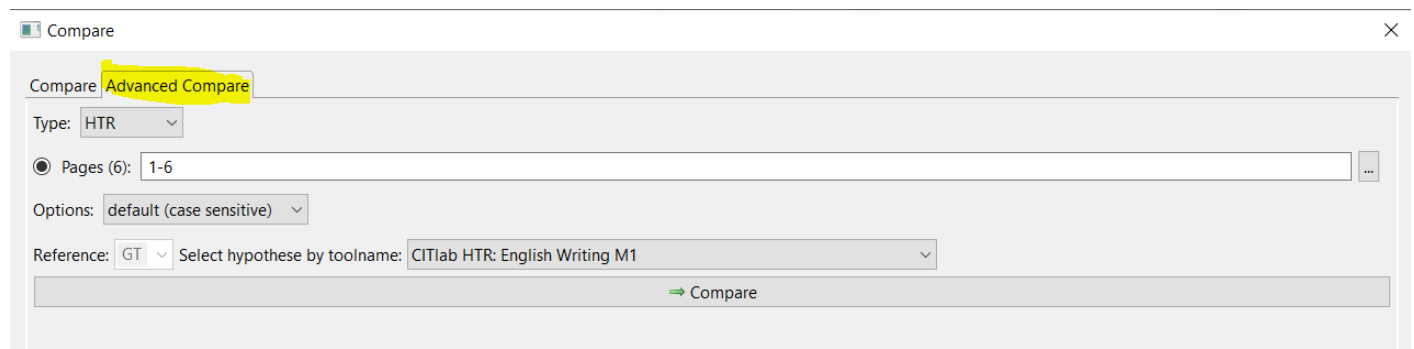
Figure 5 Advanced Statistics

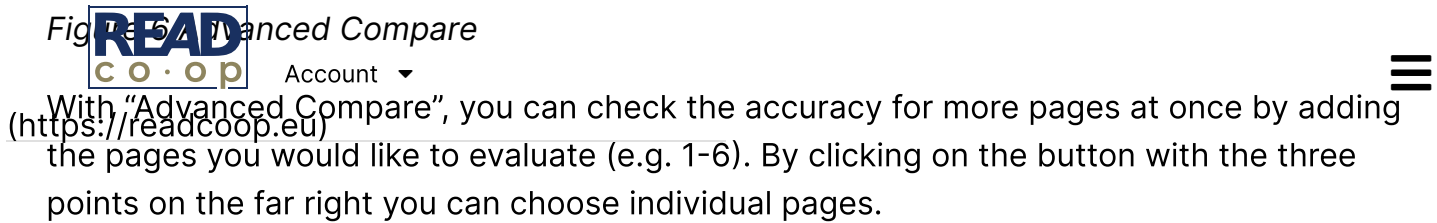
The overview display shows two tables: one with the “Overall” value, which are the average values of the recognition on all pages in a document. In the table below you can find the values for the individual pages. This way you can compare the results on different pages and by double-clicking the line you will arrive at the text-comparison, where you can check which words or text passages have been challenging.

Note: The weighting of pages for the “Overall” value is calculated based on the number of recognised words on a page.

Advanced compare

When opening the “Compare” window you can choose another tab called “Advanced Compare”.





After starting the accuracy check by clicking on "Compare", the results will be shown in the table below and by double clicking the value in the "Created" column, you will arrive at the "Advanced Statistics" window again.

Compare Samples

The "Compare Samples" functionality is useful if you are planning a bigger recognition project and would like to evaluate which model to choose before you run it on the whole document. This comparison feature chooses random lines from the sample document and tests the performance of the model on these lines.

It makes sense to put some pages aside at the beginning in order to use them as sample documents. This is advantageous, as the material the model will be tested on has not been seen before and therefore the evaluation result will be more reliable.

The "Compare Samples" functionality is also situated within the "Tools" tab in the "Compute Accuracy" section. To open it, click on "Compare Samples" and under "Create New Samples" fill out the required information.

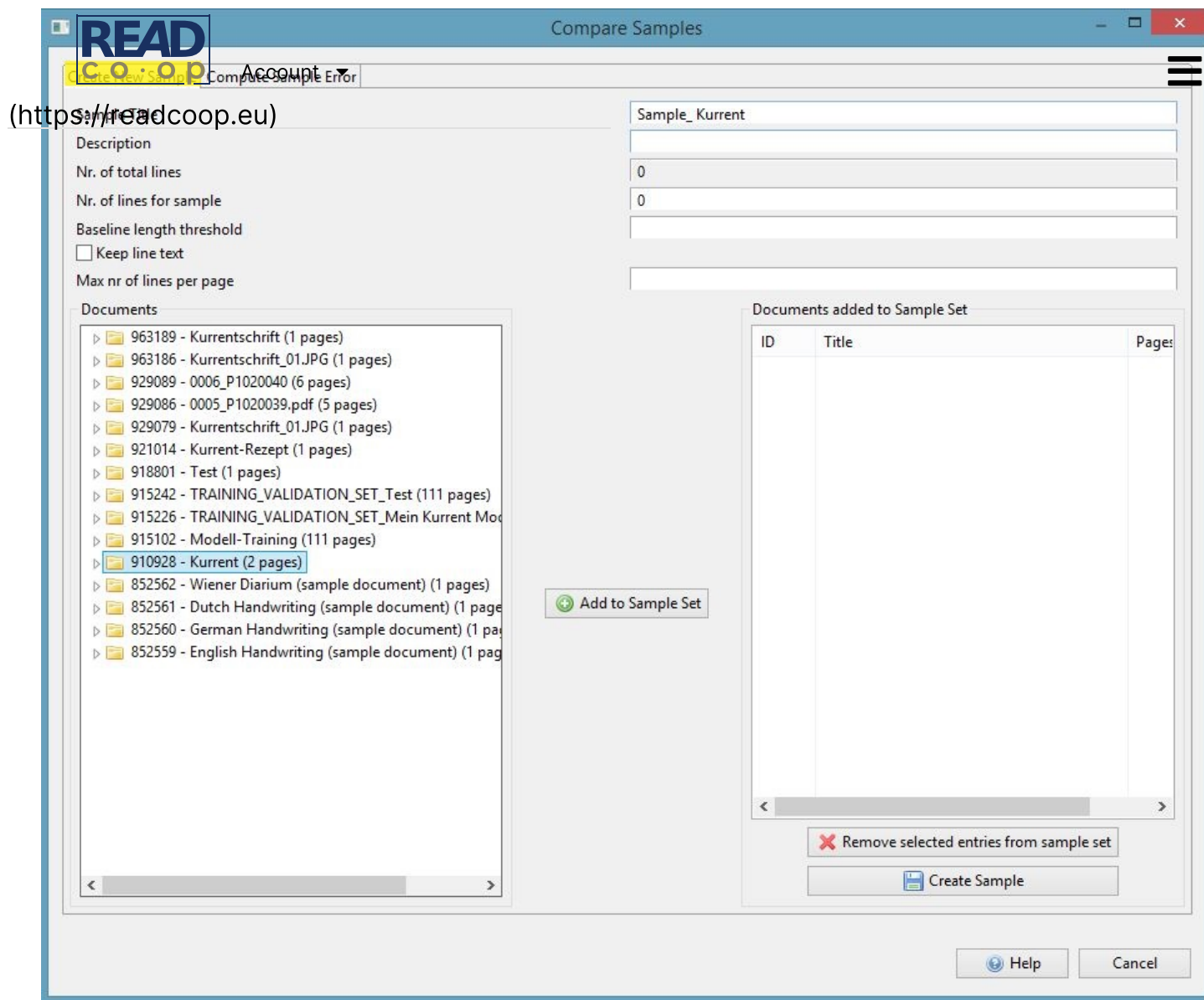


Figure 7 “Create New Samples” window

At “Nr. of lines for sample” you can define how many lines you would like to test. 500 are a recommended average. The more lines you put here the lower the variation will be in the result and the prognosis will be more precise. For a large project with many pages, it might be reasonable to say 1000 lines, for a very small attempt, maybe 100 lines are already enough. Here too, like with so many things, the best way to go about it is a “trial and error” approach, as it always depends on the individual goal.

With the “Baseline length threshold” you can control the length of lines, which is practical if you have a lot of short lines in your material, which often happens e.g. with tables in newspapers. So you can say, e.g. a line should be at least 20% of the line width – for handwritten material with only one column this step is probably not necessary. By clicking the “Keep line text” option, you can literally keep the text you already have in your documents and only need to correct the lines after creating the sample.

From the **READ** on the left side choose the collection and document of which the sample should consist of via the “Add to Sample Set button”. Then click on “Create sample”. Transkribus will now randomly choose the defined number of lines in the selected documents.



The next step is to load the sample document (you can find it in your collection) and manually transcribe the line snippets (if you haven’t kept the text like described above). It will be only one line per page and therefore the transcribing in most cases will be quick. If you have finished one line jump to the next page of the sample document to proceed.

When you are finished with transcribing, run the model which you would like to test on the sample document – to produce the transcription you can then compare with the “Compare Samples” feature.

To do so, open the “Compute Sample Error” tab in the “Compare Samples” window and choose the document you would like to evaluate. Then click on “Compute” to start the job. As soon as “Completed” appears in the “Status” column you can double click the cell in the “Created” column to view the results.

Documents **Samples**

Sample Collection

- > 141855 - mist2 (10 pages)
- > 141798 - mist (10 pages)

Reference :

GT

Select hypothesis by toolname :

CITlab HTR: WrDiarium_M8

Compute

Upper bound : 4.770%

Lower bound : 1.099%

Mean : 2.935%

With the probability of 95% the CER for the entire document will be in the interval [1.099% 4.770%] with the mean : 2.935%

By taking 4 times the number of lines the interval size can

CER per sample page

Created	Status	Queries
06.02.21 18:28:11	Completed	Page(s) : 1-10 O

Results

Created	Status	Queries
06.02.21 18:28:11	Completed	Ref: GT Hyp: CITlab

Help Cancel

Figure 8 Generating results with the “Compare Samples” feature



Account ▼


[\(https://readcoop.eu/\)](https://readcoop.eu/)

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/video/>)



Account ▼

(<https://readcoop.eu>)

Community

ps:/ link e.c /w
/tw edi ww.
tter n.c fac
om/ cha ebo
co om/ nne
m/T lco 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
jA)

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