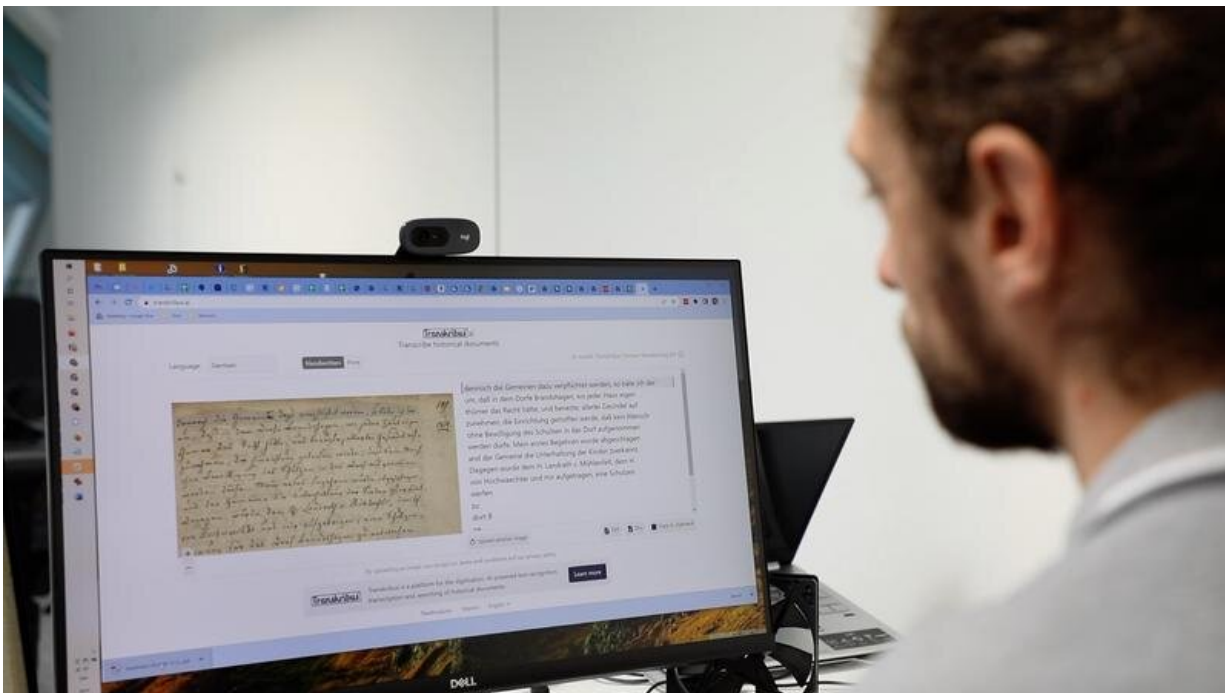


Reading old handwriting with an AI platform

October 11 2022, by Christian Flatz



Using the online platform, even non-experts can digitize and read historical manuscripts very easily. Credit: Universität Innsbruck

Using artificial intelligence, computers can decipher handwritten texts and make them readable for everyone. The Transkribus platform, co-developed at the University of Innsbruck, Austria, makes this technology available to scholars and the general public. An ever-growing group of people are using Transkribus to research their family history. Recently, users from all over the world met in Innsbruck.

Handwriting is as individual as people are. Nevertheless, computers today are capable of automatically recognizing handwritings in a wide variety of languages. The Transkribus software platform, co-developed by the University of Innsbruck, makes this technology available to the [scientific community](#), archives, and the general public. Over 90,000 users from all over the world are already using the platform to make handwritten documents readable and searchable. An ever-growing group of people are interested in their [family history](#) and are beginning to search for their ancestors in church records, contracts, or other [historical documents](#).

"Searching these documents by hand can be a very tedious task. Our technology now makes researching family history much easier," says Günter Mühlberger from the Digitization and Digital Archiving Working Group at the University of Innsbruck, Austria, and Chairman of the Board of Directors of the European cooperative READ-COOP.

Quickly search large collections

Archives and libraries store historical documents of inestimable value. These documents take up a lot of space. For example, the documents in the Austrian State Archives fill 350 shelf kilometers. Most of these documents are only available in handwritten form and are no longer legible for many users because they are written in a script called Kurrent, an old form of German-language [handwriting](#) based on late medieval cursive writing.

"This is where the Transkribus platform comes in handy, automatically recognizing this handwriting and thus making it readable for everyone," explains Günter Mühlberger. In addition, the documents can also be easily searched. This makes research using historical collections much easier because hundreds or thousands of documents can be searched simultaneously for family names or other terms.

Reading German Kurrent, Arabic, and Chinese

Transkribus works with neural networks. This machine-learning method has the great advantage that you no longer have to manually program recognition for each type of writing. "The users teach the machine to read the handwriting," says Günter Mühlberger. "And a machine does not get tired, which means it can process thousands, hundreds of thousands, or millions of pages, automatically. That's what we did for the National Archives of Finland, for example, where more than 2 million handwritten documents dating back to the 19th century are now searchable for everyone."

The technology used is completely independent of the language and the actual script or type of writing. Transkribus recognizes not only German Kurrent or modern handwriting, but also medieval scripts, as well as Hebrew, Arabic, or Indic handwriting. "And right now, we are experimenting with ancient Chinese," Mühlberger adds.

A great help for researchers

In science, and the humanities, too, the applications of Transkribus are manifold. For example, the Innsbruck classical philologist William Barton, who received the 1.2 million euro START Prize for his research with the help of Transkribus, decoded diary entries of Karl Benedikt Hase from the 19th century that were thought to be lost, penned in handwritten ancient Greek. Valuable information contained therein is to be made accessible to other fields of research.

"The private and secret diaries of the scholar Karl Benedikt Hase contain records from nine years. The amount of text is enormous, there are about 2,500 pages," William Barton of the Department of Neo-Latin Studies explains. "I trained the machine to model Hase's handwriting based on 100 pages. Now it's capable of reading all of his diaries and transcribing

the text reliably." A recent study by the University of Edinburgh revealed that more than 400 [scientific publications](#) have now been produced with the help of Transkribus.

More information: Transkribus: readcoop.eu/transkribus/

Provided by Universität Innsbruck

Citation: Reading old handwriting with an AI platform (2022, October 11) retrieved 9 April 2024 from <https://techxplore.com/news/2022-10-ai-platform.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--