

WINTERSCHOOL | NOVEMBER - DECEMBER 2024

HANDWRITTEN TEXT RECOGNITION OF MEDIEVAL DOCUMENTS

 CAROLINGIAN LATIN | LATE MEDIEVAL LATIN | BYZANTINE GREEK |
 MEDIEVAL CZECH | MEDIEVAL GERMAN | SYRIAC


"A medieval monk works on his laptop in the scriptorium." The AI image was created on 28.05.2024 by the Bing Image Creator.

ORGANISATION

Institute for Medieval Research
 of the Austrian Academy of Sciences, Austria
 Manuscript, Rare Books and Archival
 Studies Initiative (MARBAS)
 Princeton University, USA
 Faculty of Arts
 Comenius University in Bratislava, Slovakia

CONTACT

Leon Pürstinger
leon.puerstinger@oeaw.ac.at
 Jan Odstrčilík
jan.odstrcilik@oeaw.ac.at

CALL FOR APPLICATIONS
WINTER SCHOOL | NOVEMBER - DECEMBER 2024
HANDWRITTEN TEXT RECOGNITION OF MEDIEVAL DOCUMENTS

We are witnessing a groundbreaking transformation in the study of manuscripts due to machine learning tools which have been developed to enable automatic transcription of documents. These innovative tools are becoming a significant part of large and small editorial projects, and are being leveraged by libraries to enhance the accessibility and searchability of their collections.

For the third year in a row, the Institute for Medieval Research of the Austrian Academy of Sciences, in collaboration with MARBAS, Princeton University and the Faculty of Arts, Comenius University in Bratislava, will host another HTR Winter School, focusing on the automatic recognition of handwritten texts in medieval documents. The course consists of two main parts: 3 online sessions (Zoom, Friday afternoons) and a three-day in-person workshop in Vienna. Between meetings, we will use two online platforms, Notion.so and Discord.com, to coordinate our work. During the first phase, participants will be briefly introduced to the theory of handwritten text recognition and comprehensively instructed in its practical application using the tool Transkribus (transkribus.eu).

We will work in **six groups**, focusing on various languages and scripts:

CAROLINGIAN LATIN
LATE MEDIEVAL LATIN
BYZANTINE GREEK
MEDIEVAL CZECH
MEDIEVAL GERMAN and
SYRIAC

Each group will have its own supervisors and will transcribe passages from selected manuscripts. We will then use these passages for training new HTR models and improving existing models, while learning about various intricacies and challenges. Participants are also welcome to consult their own material in addition to the team work.

During the in-person workshop in Vienna, we will finalise the projects and visit libraries in Vienna to view the selected manuscripts. At the end of the course you will receive a certificate and the results will be published on Zenodo and HTR-United with the names of the contributors.

REQUIREMENTS

The course is primarily designed for Master's or PhD students, however, we will consider other applicants as well.

You are expected to be familiar with the language of the group you want to join: Carolingian Latin, Late Medieval Latin, Byzantine Greek, Medieval Czech, Medieval German, and Syriac.

We expect you to have knowledge of medieval palaeography and manuscript studies. If in doubt, don't hesitate to ask.

The course will be taught in English.

COSTS

There is no participation fee, but you are expected to cover the cost of your travel to Vienna, including accommodation.

APPLICATION

Please, send a short CV and cover letter to HTR.school@oeaw.ac.at and to Leon.Puerstinger@oeaw.ac.at before **September 15th 2024** with "HTR School 2024" as the subject. Also, please don't

forget to indicate the team/module you would like to take (i.e. **Carolingian Latin, Late Medieval Latin, Byzantine Greek, Medieval Czech, Medieval German, Syriac**). If you have any questions, don't hesitate to ask.

ORGANISERS (alphabetically)

Eirini Afentoulidou | Austrian Academy of Sciences
Eirini.Afentoulidou@oeaw.ac.at

Cinzia Grifoni | Austrian Academy of Sciences
Cinzia.Grifoni@oeaw.ac.at

Ephrem Aboud Ishac | Austrian Academy of Sciences
aboud.ishac@oeaw.ac.at

Ivana Lukáč Labancová | Comenius University in Bratislava | ivana.lukac@uniba.sk

Daniela Mairhofer | Princeton University
daniela.mairhofer@princeton.edu

Anna Michalcová | Czech Academy of Sciences
a.michalcova@ujc.cas.cz

Ekaterini Mitsiou | University of Vienna
ekaterini.mitsiou@univie.ac.at

Jan Odstrčilík | Austrian Academy of Sciences
jan.odstrcilik@oeaw.ac.at

Leon Pürstinger | Austrian Academy of Sciences
leon.puerstinger@oeaw.ac.at

Helmut Reimitz | Princeton University
hreimitz@princeton.edu

Martin Roček | Charles University in Prague
martin.rocek@ff.cuni.cz

Christine Roughtan | Princeton University
croughan@princeton.edu

Gerda Rummel-Heydemann | Freie Universität Berlin | gerda.rummel-heydemann@fu-berlin.de

William Weis | University of Vienna
Bill.Weis@oeaw.ac.at

Michaela Wiesinger | Austrian Academy of Sciences
michaela.wiesinger@oeaw.ac.at



SCHEDULE

3 ZOOM ONLINE SESSIONS

1 Friday NOV 8	What is HTR? A general introduction
	Transkribus 1 (uploading documents, layout recognition, simple transcription)
	Introduction into manuscripts and working in groups
2 Friday NOV 22	Transkribus 2 (introduction to models, CER, learning curve)
	Working in groups (first transcriptions)
3 Friday DEC 6	Exporting documents / Sharing is caring: How to share and publish your data 1 (introduction)
	Tagging structures and text
	Working in groups (training of models)

3-DAY-WORKSHOP IN PERSON AT VIENNA

Wednesday DEC 18	Finishing work on the transcription and models
	Sharing is caring: How to share and publish your data 2
	Working in groups
Thursday DEC 19	Alternatives to Transkribus
	What can you do with automatic transcriptions?
	Library visit - See the manuscripts
	Working in groups
Friday DEC 20	Time to finish the work
	Presentation of results by working groups