Workshop A: Anissava Miltenova, Andrej Bojadžiev (Sofia): Electronic Descriptions of Slavic Manuscripts (Repertorium Initiative)

The workshop provides an introduction to methods and technology used in the framework of Repertorium Initiative (http://repertorium.obdurodon.org/) – a multinational project for electronic description of Medieval and Early Modern Slavic manuscripts. Some topics will include: levels of representation of manuscript descriptions; what is so special about the Repertorium model; how to provide information on codicological, palaeographic and language characteristics of manuscripts; how to encode the manuscript content with special emphasis on text history and the place of text witness among the tradition; how to edit colophons and scribal notes. We will discuss some topics like how to markup various names in order to establish authorative indices and how to represent our electronic description.

Prerequisites: The participants are requested to download Oxygen editor at https://www.oxygenxml.com/xml_editor/download_oxygenxml_editor.html). Clicking on download link activates a download and enters registration page. Please fill in all fields, choose at License Type "Academic". Then you can request a 30-day evaluation key. After the installation you will be asked for this key which you will obtain per email.

Workshop B: Yves Scherrer (Helsinki): Neural tagging of historical language varieties

In this workshop, we will discuss recent tagging techniques based on neural networks and see how such techniques can be adapted and applied to historical data. Based on our experience with tagging of pre-modern Orthodox Slavic varieties, the workshop will start with an overview of different tagging tasks (Part-of-speech tagging and full morphological tagging) and different recent technical approaches to tagging. We will also discuss the specific challenges of historical data such as the high orthographic variability and the generally small size of available data. The workshop will provide ample space for hands-on experiments with a neural tagger, where participants will learn how to train neural tagging models and use them to annotate new data.

Prerequisites: The hands-on experiments will require access to a working Python 3 installation and basic knowledge of command-line tools. Programming knowledge is not required. Further details about the technical setup will be communicated in the first workshop session.

Workshop C: Achim Rabus (Freiburg): Introduction to Transkribus (not only) for Slavists

In this workshop, the participants get a hands-on introduction to Transkribus (www.transkribus.eu), arguably the most advanced AI-assisted Handwritten Text Recognition (HTR) application currently available. After an overview of the basic principles of HTR and their implementation in Transkribus (layout analysis, model training), the capabilities and limits of specific HTR models relevant (not only) to Slavists will be discussed. The participants learn how to transcribe their own documents using models already publicly available and how to train their own models if there are no suitable public models available.

Prerequisites: The participants are requested to create a Transkribus account, to download the Transkribus desktop application and to register at www.transkribus.eu. After successful registration, they are asked to fill out an online questionnaire no later than April 5. We will create a joint Transkribus collection for the workshop where every participant can upload parts of the sources they are interested in transcribing using Transkribus.