Unlocking the Historical Ledgers of the Bank of Amsterdam (MultiX Lab) n.j.e.vannoord@uva.nl

Description
The ledgers of the Bank of Amsterdam (Amsterdamse Wisselbank) from between 1650-1800 contain a wealth of information. Yet, the information in these approximately half a million pages is all hand-written and not easily extracted. Researchers of the GLOBALISE project (https://globalise.huygens.knaw.nl/) are using Handwritten Text Recognition (HTR) for parsing these ledgers, to try and unlock the information within them. As digitising these ledgers would enable large-scale research into the financial traffic within Amsterdam and the historical role of the bank.

Usually, HTR is applied to letters and written texts allowing for straightforward line-by-line parsing - the ledgers, however, follow a tabular format which is more structured, but deviates from what HTR approaches expect. In this project we would like to explore how the process of extracting handwritten text from historical tabular documents can be improved. Possible solution directions include document parsing, novel methods for preprocessing, and/or black-box optimisation.

Work Environment
Supervised by Nanne van Noord of the MultiX lab and in collaboration with Lodewijk Petram (Huygens KNAW).

Expectations
Experience with Computer vision and an interest in historical research is a plus for this project.