

MELA DSIG Virtual roundtable on Arabic script OCR / HTR

In the midst of a host of digitization projects for enhanced text access, developments in Arabic script automated text recognition (OCR / HTR) are taking place rapidly and with increasing momentum.

As collection workers & information specialists documenting, preserving, and enabling access to text based materials in Arabic script, our community is invested and engaged in various ways with this work.

To better acquaint our colleagues with recent developments, the [Middle East Librarians Association Digital Scholarship Interest Group \(MELA DSIG\)](#) is convening a roundtable discussion showcasing the experiences of a sampling of colleagues working with various Arabic script OCR / HTR tools, workflows, materials, and languages. With this roundtable, we hope to offer reflections on progress with Arabic script OCR / HTR performance, costs and other factors as well as to invite our community to consider the implications of this advancing technology for collecting, preservation, and research and our ongoing contributions to this work as library & archives specialists.

Thursday, 1 October 2020 11-1 EDT via Zoom (link will be provided following [registration & payment for the MELA Conference](#))

Introduction (10 minutes)

Evyn Kropf (University of Michigan)

Roundtable (50 minutes)

Discussant: Dale J. Correa (University of Texas Austin)

Akram Khater (North Carolina State University Khayrallah Center)
Tesseract adapted for [Arab American Newspapers Project](#)

Peter Magierski (Columbia University)
Work with NovoVerus

Suphan Kirmizialtin (NYU Abu Dhabi)
Transkribus HTR for automated transcription of printed Ottoman Turkish
(cf [Transkribus in practise talk](#))

Moderated discussion / Q&A (20 minutes)

Report from the [Open Islamicate Texts Initiative Arabic-script OCR Catalyst Project \(OpenITI AOCP\)](#)

Matthew Thomas Miller (University of Maryland) (15 minutes)

Concluding discussion / Q&A (20 minutes)

This event is being held in conjunction with the [MELA 2020 Conference](#)