3rd summer school in
Islamicate digital humanities
for scholars in manuscript studies

‹ﺽ/›
islamicate_DH
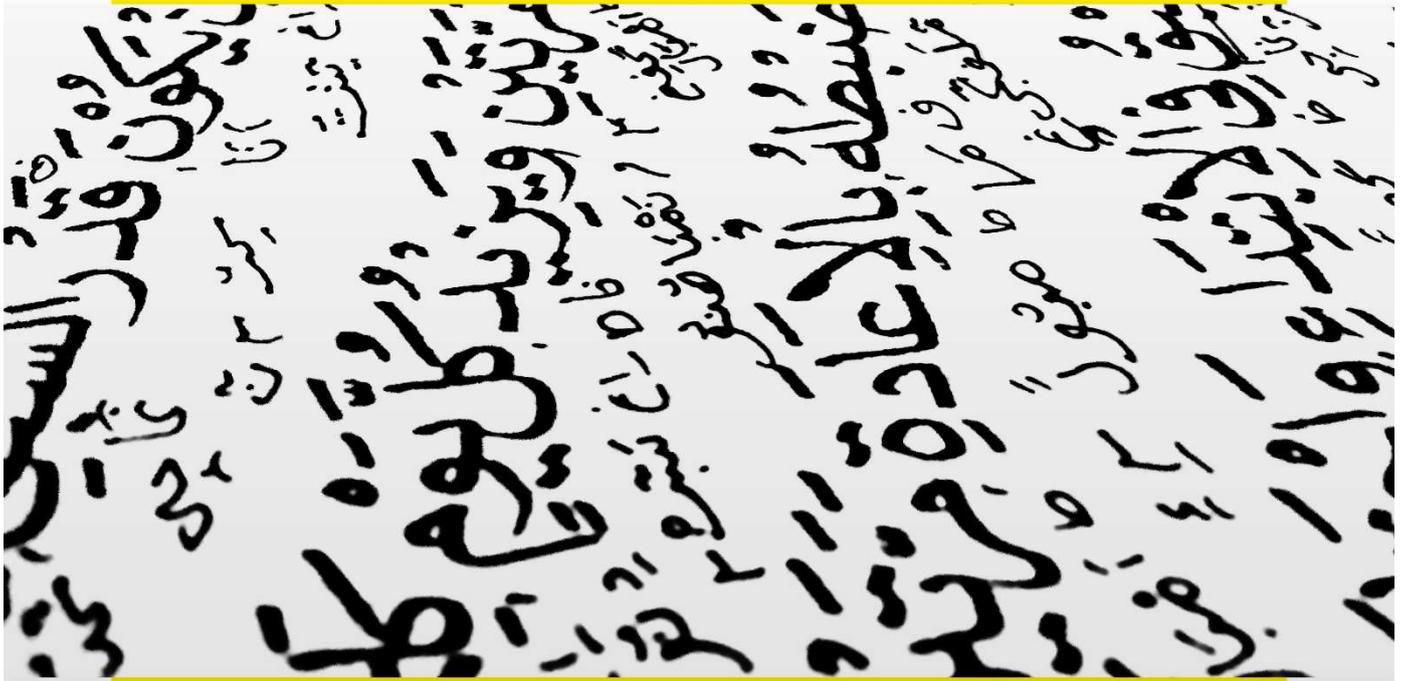
Date: 30 August – 10 September 2021 (twice a week, on Mondays and Fridays)

Venue: online, hosted by Hamburg Universität, Edmund-Siemers-Allee 1, 20146 Hamburg

Director of the Course: Alba Fedeli

IT Coordinator: Alicia González Martínez and Cristina Vertan

Deadline for applications: 2 July 2021

Registration free of charge

Number of Participants: a limited number of participants will be selected, up to 15 people

Contact details: Alba Fedeli at alba.fedeli@uni-hamburg.de

Aims and content

This summer course is for graduate (MA and PhD) students and researchers who work with manuscript materials in Arabic script and want to learn how to choose tools and resources that are available to them and how to apply digital technology in their research.

The course will consist of introductory sessions on basic topics scholars encounter when dealing with manuscripts in Arabic script in a digital environment: text encoding, selection of an appropriate model for the organisation of the data and for annotation, mark-up languages and tools, surface- and semantic based search. This theoretical framework will be followed by hands-on sessions.

Instructors will guide participants in a practical session for planning their own digital edition of a few manuscript leaves using TEI standards and, optionally, their own manuscript texts.

There will be practical sessions for learning how to perform advanced text search using regular expressions, and understanding the added-on value of a semantic search and which digital resources have to be involved in such a search.

At the end of the course, participants will be able to plan and build a digital edition of their manuscript texts and will have a basic knowledge of the powerful tools that they have at their disposal for making the most of their data to investigate manuscripts. Ultimate goal of the summer school is to give a basic knowledge on how to choose – if a digital tool is needed - among the numberless options available to scholars according to their own research questions.

Maxim Romanov will give a keynote lecture on "Prospects of Middle Eastern Studies in the context of Digital Humanities".

Practical information

- The number of participants is limited to **fifteen people** for the effectiveness of the practical sessions and thus, people who intend to participate must apply **by 2 July 2021**.

- The registration - free of charge – is mandatory for the selected participants as the number of available places is very limited.

- The language of instruction will be **English** and participants are required to have knowledge of **Arabic script**.

Application

As the number of participants is limited to fifteen people for the effectiveness of the practical sessions, students and researchers interested to participate in this summer course must apply by submitting the following details in a single PDF-file:

- A **one-page** CV
- A motivation letter (**maximum 200 words**) explaining their interest in the course and their field of research
- A **list of their technical skills** in order to organize the practical sessions in the more effective way although they are not a requirement for participating in the course. Applicants have to indicate – for example - their knowledge of -any software for annotation, searching and/or any type of data handling; data formats; digital editing; programming languages, etc.

All documents have to be sent as a single PDF-file attachment named as "LastName_Name_Hamburg_DHSchool_2021" to Alba Fedeli at alba.fedeli@uni-hamburg.de and Alicia González Martínez at alicia.gonzalez@uni-hamburg.de

The deadline for applications is **Friday 2 July 2021** and participants will be informed about their application by **Friday 16 July 2021**.

Please, consider that the time indicated is Central European Summer Time (i.e. Hamburg Time)

The schedule could be adjusted according to the applicants who will be selected in order to facilitate the attendance in their time zone

30 August 2021 Monday	3 September 2021 Friday	6 September 2021 Monday	10 September 2021 Friday
technical checks			
Welcome			
Introduction to DH, Overview, Corpus Based Research	RegEx Theory&Practice	TEI Practical Part	RegEx Theory&Practice
		Data Conversion	
Data Formats	InterSaME		
TEI Theoretical Part	Ontologies Theoretical Part	Ontologies Practical Part	Keynote Lecture Maxim Romanov
			Conclusions

h.9.00-9.30
Warm-up and technical checks

h.9:30-10:00
Alba Fedeli
Welcome and introduction of instructors and participants

h.10:00-12:00
Alicia González Martínez and Cristina Vertan
Introduction to digital humanities and corpus-based research
      Theoretical orientation
The instructors will show how humanities can benefit and make use of digital technologies and provide a broad overview of state-of-the-art methods, tools and resources in DH in general and the particular challenges for Arabic-scripted texts.

h.13:30-14:30
Alicia González Martínez
Encoding and data formats

In the first part of this session participants will understand how computers represent script and how to choose the right encoding for their work. In the second part of the session participants will be able to distinguish between unstructured and structured digital data and which are the most relevant data formats for their work.

h.15:00-18:00
Florian Jäckel
Digital transcriptions and editing (TEI)
      Theoretical orientation and exercises together, guided by the instructor
Building upon the introduction to Data Formats, the first session will offer a theoretical orientation to TEI and to key aspects of TEI markup for digital editions, including meta-data in headers, markup of a text's structure as well as recording variants.

h.10:00-12:00

Alicia González Martínez

Regular Expressions

Theory and practice

Regular expressions are the best device for doing powerful searches on raw texts. The instructor will show how participants can use them and a general overview of common software that includes them.

h.13.00-14.00

Alba Fedeli and Alicia González Martínez

InterSaME: Annotating and structuring your edition and manuscript images

During this short session, the instructors will show a software tool to annotate manuscript images and edit their texts. The tool answers specific needs and research questions but – at the same time - generates new approaches for looking at manuscripts. The research and tool are part of the DFG-AHRC InterSaME project (https://www.intersame.uni-hamburg.de/project.html).

h.15:00-17:00

Cristina Vertan

Building a knowledge base for semantic annotation and search

Theoretical orientation

Research in Humanities, in particular in manuscript studies relies to a great extent on background domain knowledge. Therefore any digital tools for qualitative (and not just quantitative/ statistical) analysis should be able to embed such knowledge and allow semantic search. In this session, participants will learn how to develop and manipulate such a knowledge base (ontology) and learn the basics of semantic search.

h.10:00-12:00

Florian Jäckel

Digital transcriptions and editing (TEI)

Exercises (with TEI edition proposed after the autonomous work done by the participants)

In this second TEI session, participants will gain practical experience by working with and expanding upon the examples presented by the instructors.

h.13.30-14.30

Carolin Kinne-Wall and Florian Jäckel

Data Conversion

Introduction and Exercises

This session provides a brief introduction to the concept of data formatting. The focus lies on converting a TEI text (using the same examples as in the foregoing session on TEI) to a well readable html webpage, using XSLT.

h.15:30-17:30

Cristina Vertan

Building a knowledge base for semantic annotation and search

Practical session

In this session, participants will have the opportunity to develop small components of an ontology, and practice queries on a larger ontology modelling the Ottoman world.

h.10:00-12:00
Alicia González Martínez
Regular Expressions
> Theory and practice

h. 14:00-16:00
Maxim Romanov
Prospects of Middle Eastern Studies in the context of Digital Humanities

> Keynote lecture followed by Q&A

Abstract: Digital humanities has become a prominent part of academic life in the course of the past two decades. Nonetheless, even digital humanists do not agree among themselves when it comes to defining this new academic domain. The lecture will offer a vision of digital humanities that aims at adapting Middle Eastern studies to the digital realities of the rapidly developing world and proposes a novel approach to the study of the Near and Middle East. At the core of this approach lies the idea of "associative trails" formulated by the American scientist and engineer Vannevar Bush (d. 1974) in the middle of the last century. Generated algorithmically, these trails would bring all available information into networks of new knowledge. In their turn, these networks will become both a terra nova of sorts that researchers can travel through for the close reading of historical evidence and new macro objects that researchers would engage with to do the distant reading of historical evidence. The lecture will highlight early examples of such networks generated from modern research publications and premodern Arabic textual sources.

h.17:00-18:00
Alba Fedeli
> *Final Remarks*

Alba Fedeli is a research associate at the Asien-Afrika-Institut, Universität Hamburg working on the transmission of early Qur'ānic manuscripts and their writing system. She received her PhD from the University of Birmingham, UK, at the Institute for Textual Scholarship and Electronic Editing, after studies in Italy with Sergio Noja Noseda and has first-hand knowledge of manuscript collections scattered all over the world, from Yemen to Ireland, from Russia to Qatar and Egypt. Her works include the edition of the Mingana-Lewis Qur'ānic palimpsest (https://cudl.lib.cam.ac.uk/view/MS-OR-01287-SMALL/1). She is the co-PI of InterSaME Project ('The intertwined world of the oral and written transmission of sacred traditions in the Middle East'), a DFG-AHRC joint project in co-operation with Geoffrey Khan at the University of Cambridge.

Alicia González Martínez is a research associate at the Asien-Afrika-Institut, Universität Hamburg. She has a multidisciplinary profile covering Arabic linguistics, Oriental studies and software development, with a special focus on Arabic Natural language processing. She worked in COBHUNI project from 2015 to 2020 and now, she is a researcher in the InterSaME project, with Alba Fedeli.

Florian Jäckel studied Catholic theology and Islamic studies at Tübingen. In 2021, he submitted his Ph.D. thesis on conceptions of prenatal life in the writings of the Syriac polymath Barhebraeus (1226-1286) at the University of Hamburg. His research interests include digital means in historical and philological research, the intellectual history of Christianity and Islam and cross communal interaction in the medieval Middle East.

Carolin Kinne-Wall works as IT-coordinator at the Asien-Afrika-Institut, Universität Hamburg. She studied Computer Science in Bonn, Cologne and Singapore and worked in different interdisciplinary projects as IT- and data manager.

Maxim Romanov is a Universitätsassistent für Digital Humanities at the Department of History, University of Vienna, and a senior research fellow at the KITAB Project led by Sarah Savant at the Aga Khan University, London (http://kitab-project.org). Maxim Romanov plays leading roles in several international projects that focus on the digitization of Near and Middle Eastern studies: 1) the development and curation of the digital corpus of Islamicate texts (OpenITI, https://github.com/OpenITI); 2) modeling of the Arabic written tradition through a flexible method of tracing text reuse (KITAB Project); development of the geospatial model of the medieval Islamic world (https://althurayya.github.io/); optical recognition of the

printed and written Arabographic text (OpenITI AOCP). In the coming months, Maxim Romanov will be joining the University of Hamburg as the junior research group leader (Nachwuchsgruppenleiter) of the project "The Evolution of Islamic Societies (c. 600-1600 CE): Algorithmic Analysis into Social History", funded by the German Research Foundation (DFG) within the framework of the Emmy Noether Programme.

Cristina Vertan is senior researcher at the Berlin Academy of Sciences and associated researcher at the University of Hamburg. She studied Computer Science in Bucharest and Marseille as well as statistics in Brussels. She holds a Ph.D. in Computer Science from the University of Bucharest and had a Humboldt fellowship in natural language processing. She is specialised in Computational Linguistics, with focus on multilingual applications and Digital Humanities, especially the processing of historical texts. She developed the first annotation tool for classical Ethiopic, and classical Maya which enables multi-level mark-up and a deep automatic controlled annotation of this Semitic language. Currently she completed the development of the first comprehensive ontology for the history of the Ottoman empire.