Digital Humanities
UNIVERSITÄT LEIPZIG

الرقميات العربية والفارسية

فارسی و عربی دیجیتال

Digital Arabic & Digital Persian

Research Workshop

Dates: Thursday–Friday, December 17–18, 2015
Place: University of Leipzig, The Humboldt Chair for DH

Description: The workshop seeks to bring together scholars of Arabic and Persian textual traditions (broadly defined), including both those who are already using digital methods and those who want start using them. The workshop will allow us to share our findings, get feedback on work in progress, and chart future research projects. The event will also give us all an opportunity to work out collaborative strategies that will be beneficial for all involved in the DH-ization of Arabic and Persian studies.

Newcomers to the digital world will have an opportunity—and are even encouraged—to partake in the hands-on methodological workshop (Dec. 14–16) that will take place before the research workshop. The methodological workshop will cover such topics as building a corpus, creating training data, topic modeling, text reuse, georeferencing, mapping, social network analysis, stemmata, and semantic analysis.

Participants will present their work—completed, in progress, planned, contemplated—and will have an opportunity to receive feedback from colleagues with more experience working with historical languages.

Framework: The 2-day workshop will include over 25 speakers. All participants are strongly encouraged to discuss collaborative efforts, with the workshop ending with a charted course for action.

Applications: If you are interested in participating, please email maxim.romanov@uni-leipzig.de. Although the focus of the workshop is on Arabic and Persian written sources (600–1500 CE), scholars studying texts from other periods and in other Islamicate languages are welcome to apply.

Working Schedule

	Thursday, December 17 (Augustusplatz 10, Neu Paulinum P702)		
1	9:00	10:25	***Toward Collaborative Future***
		1	Maxim Romanov, U Leipzig—Introduction: Building the Future Together
		2	Ouafae Nahli, Istituto di Linguistica Computazionale "A. Zampolli" (ILC) of the National Research Council (CNR), Pisa—*Improved Written Arabic Word Parsing through Orthographic, Syntactic and Semantic constraints*
		3	Till Grallert, Orient Institut Beirut—*The journal* al-Muqtabas *between Shamela.ws, HathiTrust, and GitHub: producing open, collaborative, and fully-referenceable digital editions of early Arabic periodicals—with almost no funds*
2	10:30	11:55	***Building Foundations: [Diachronic] Corpora***
		4	Matthew Miller, U Maryland—*Workshop on the Persian Digital Library (PDL): Status Update and Strategic Planning*
		5	**SKYPE** Sébastien Moureau, the Warburg Institute (London)—*Qawl 1.0 & 2.0: intentions, actual state, and future development*
		6	Peter Verkinderen, U Hamburg & Maxim Romanov, U Leipzig—*Building a Unified Arabic Corpus: Open Discussion*
3	12:00	1:25	***All we need is... OCR***
		7	Federico Boschetti (U Pisa), Arslane Zarghili & Younes Lasri (Sidi Mohamed Ben Abdellah U, Morocco)—*Enhancing the Arabic OCR workflow*
	1:30	**3:00**	**Lunch Break**
4	3:00	4:25	***Text as big data***
		8	Sarah Savant, Agha Khan U, London—*KITAB-Project: Tracing Text Reuse in the Arabic Tradition*
		9	Jiří Milička & Petr Zemánek, Charles U, Prague—*Tolerant algorithm for quotation extraction and its application on diachronic Arabic corpus*
		10	Matthew Miller, U Maryland—*Computational Explorations of Genre in 'Classical' Persian Poetry*
5	4:30	6:00	***Persian & Urdu***
		11	**SKYPE** A. Sean Pue, Michigan State U—*Topic Modeling Urdu Poetry*
		12	**SKYPE** Mushegh Asatryan, U Calgary—*Iranian Loanwords in Classical Arabic: Language Contact and Sound Change*
		13	Omid Azadibougar, U Göttingen—*World Literature and English Translations of Persian Poems*
	Friday, December 18 (Augustusplatz 10, Neu Paulinum P502)		
6	9:00	10:25	***The Nitty-Gritty: Syntax, Morphology, Lexicography, Semantics***
		14	Omid Tabibzadeh, Bu-Ali Sina University, Iran—*The Autonomous Phrases Theory; A Dependency Based Approach to Syntax Processing: A Case Study on Persian Simple Sentences*

			15	Navid Hemati, Bu-Ali Sina University, Iran/University of Uppsala—*Towards a Syntactic Dependency Treebank for Classical Persian Based on the Autonomous Phrases Theory*
			16	Kirill Dmitriev, U St Andrews—*Analytical Database of Arabic Poetry (Arab Cultural Semantics in Transition)*
7	10:30	11:55		***Big Library Data***
			17	Daniel Lowe, British Library—*Digitizing Arabic Collections*
			18	Nur Sobers-Khan, British Library—*Bengali Books: Digitisation of South Asian Early Printed Material at the British Library*
			19	Daniel Brenn, Orient Institute Beirut—*Source Companion on Islamic Political Thought (SCRIPT)*
8	12:00	1:25		***Antiquity***
			20	Usama Ghad, Institute for Papyrology (Heidelberg, Germany)—*Graeco-Arabico-Latinum Encoded Corpus (GALEN©)*
			21	Elaine van Dalen, U Manchester—*Arabic Commentaries on the Hippocratic Aphorisms*
			22	Corneliu Arsene *et al.*, U Manchester/U Exeter—Uncovering the Syriac Galen Palimpsest by Using Image Processing Techniques
	1:30	**3:00**		**Lunch Break**
9	3:00	4:25		***Early Islamic History: Stepping into the Digital***
			23	Ed Hayes, Colby College—*Early Shiʻi biographical database*
			24	Majied Robinson, Edinburgh U—*Early Arab Genealogies: Datasets, methods and research questions*
			25	Yoones Dehghani Farsani, Georg-August-U in Göttingen—*The Early Futuh Tradition*
10	4:30	5:40		***Projects***
			26	Verena Klemm, Universität Leipzig—*Bibliotheca Arabica*
			27	Beatrice Gründler, Freie Universität Berlin—*Kalila and Dimna: Wisdom Encoded*
11	5:40	6:30	27	***Charting the Course: Final Discussion***