

Schedule for Baladaria Summer School in Zadar

updated: 7/6/2023

Start		Monday, June 19	Tuesday, June 20	Wednesday, June 21	Thursday, June 22	Friday, June 23
8:30	Topic	Structured data	Corpus Linguistics	Enrichment	Group work	Further resources
	Tool	<i>GSheets/Excel</i>	<i>AntConc & TagAnt</i>	<i>Recogito</i>	<i>(at CodeHub)</i>	<i>QGIS, OpenRefine</i>
	Techniques	Finding structured data; importing different file formats; basic visualizations; API; encoding	Text input; text preprocessing; basic text analysis (frequency, co-occurrence); cocordances; collocations	XML; Scholarly annotation; NER; text-to-map		spatial mapping; cleaning metadata; regular expressions
	Lead	Daniel Ocic Ihrmark, Ahmad Kamal	Daniel Ocic Ihrmark	Daniel Ocic Ihrmark		Ahmad Kamal
	Support	<i>Koraljka Golub, Marijana Tomic</i>	<i>Ahmad Kamal, Koraljka Golub, Marijana Tomic</i>	<i>Ahmad Kamal, Cécile Chantraine Braillon, Fatiha Idmhand, Koraljka Golub, Marijana Tomic</i>		<i>Koraljka Golub, Marijana Tomic</i>
10:15	<i>break</i>					
10:30	Tool	<i>Gephi</i>	<i>AntConc</i>	<i>Transkribus, AV annotation</i>	Group work	<i>Voyant Tools, Omeka, CLARIN, Tapor3</i>
	Techniques	network analysis	further corpus analysis	crowdsourcing, audio-visual annotation	<i>(at CodeHub)</i>	text analysis; digital collection building
	Lead	Ahmad Kamal	Daniel Ocic Ihrmark	Marijana, Cécile Chantraine Braillon, Fatiha Idmhand		Ahmad Kamal
	Support	<i>Daniel Ocic Ihrmark, Koraljka Golub, Marijana Tomic</i>	<i>Ahmad Kamal, Koraljka Golub, Marijana Tomic</i>	<i>Ahmad Kamal, Koraljka Golub</i>		<i>Koraljka Golub, Marijana Tomic</i>
12:00	<i>lunch</i>					
13:00	Topic	Data science	Text analysis I	Text analysis II	Group work	Ad. Processing
	Tool	<i>Python (os library)</i>	<i>Python (NLKT/Spacy)</i>	<i>Python</i>	<i>(at CodeHub)</i>	<i>Python</i>
	Techniques	Readings/writing files; web scraping (html); reading spreadsheet (CSV); importing text (txt)	Text scrapping; text preprocessing; basic text analysis	NER; Entity linking		Advanced AI (Dall-e, ChatGPT)
	Lead	Axel Jean-Caurant, Benedikt Perak	Axel Jean-Caurant, Benedikt Perak	Axel Jean-Caurant, Benedikt Perak		Dan Kohen, Benedikt Perak
	Support	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>		<i>Axel Jean-Caurant, Benedikt Perak, Marcelo Milrad, Željka Tomasović</i>
14:30	<i>break</i>					
14:45	Tool	<i>Python (Pandas, Mathlab)</i>	<i>Python (NLKT/Spacy)</i>	<i>Python (Spacy)</i>	Group work	<i>Python</i>
	Techniques	data cleaning; data frames; basic data visualization	basic text analysis; sentiment analysis; topic modeling	Regular expression; text categorization	<i>(at CodeHub)</i>	Image processing
	Lead	Axel Jean-Caurant, Benedikt Perak	Axel Jean-Caurant, Benedikt Perak	Axel Jean-Caurant, Benedikt Perak		Axel Jean-Caurant
	Support	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>	<i>Antoine Doucet, Dan Kohen, Marcelo Milrad, Željka Tomasović</i>		<i>Dan Kohen, Marcelo Milrad, Željka Tomasović</i>
16:15	<i>free time</i>					
16:30	<i>end</i>					