**Arabic Tagging Project**

8th meeting – 06.05.2020

**What we did**

MongoDB – we took Fozi’s DB it’s a relational DB with more than 2 million terms and each term associated to a poem. We took his database and changed it so that we save each poem as a file alone with an id and an array of [line, sader, ajuz].

We don’t have to save each term alone when saving them poem since python has libraries that can cut a string to terms and we’d rather do that while processing the poem and then saving each term and its tag.

Same as with poems we took the poets from Fozi’s DB and associated each poem with a poet.

Neo4j – we saved the tags hierarchy we have in a neo4j graph, it allows us to search, add, remove, view its nodes in an easy and efficient way.

**For next time**

We’re taking a break from frontend and now concentrating on the backend part, we want to process each line by cutting it, processing them term (root, حركات) and knowing which tags to suggest to the user.