

# Arabic Text Readability Analysis

Mentors: Hamdy Mubarak, Ahmed Abdelali, Kareem Darwish {hmubarak, aabeldelali, kdarwish}@hbku.edu.qa

**Motivation:** Readability scores are a way to measure whether a written text is likely to be understood by the intended reader. Text that scores a readability grade level of 8 or below will be readable by around 85% of the general public. This can help in improving language skills for **language learners**, **proficiency tests**, **text simplification**, designing of **educational materials**, etc.

**Available Resources:** Arabic **curricula** for Gulf countries (Grades 1 to 6), Automatic processing (Farasa NLP tools) for **diacritization**, **stemming**, **parsing**, etc. huge Arabic corpora from **Wikipedia**, **Aljazeera** and other sources.

We already built an initial interface for spotting **complex words** that didn't appear in a selected grade.

- ❑ We will study different **readability formulas** and apply to Arabic.
- ❑ We will build a **website** to provide **readability analysis** for an input text and give tips to improve its readability.

## ❖ Skills

- ✓ Experience in Java/Python for data processing, calling APIs...
- ✓ Experience in web development: HTML, CSS
- ✓ Passionate about language analysis and education
- ✓ Understanding Arabic is a plus!

Enter a text:

خبراء: الأثرياء أكثر جرأة في المغامرة والسفر بدأ الأغنياء القيام برحلات تحويلية تستند إلى فلسفة اكتشاف الذات. وفي الوقت نفسه، يؤكد الخبراء أنه كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً.

Grade: Grade 3

SUBMIT

Result:

خبراء: الأثرياء أكثر جرأة في المغامرة والسفر بدأ الأغنياء القيام برحلات تحويلية تستند إلى فلسفة اكتشاف الذات. وفي الوقت نفسه، يؤكد الخبراء أنه كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً، كلما كان الشخص أكثر ثراءً.

## Readability Analysis

FAVES	GRADE	ISSUES	REACH	WORDS
☆	B	20	100%	304
Readability Grade Levels				
😊	Flesch-Kincaid Grade Level	7.9	★	
😞	Gunning Fog Index	10.3	★	
Readability Scores				
😊	Flesch Reading Ease	66.8	★	
Text Statistics				
	Word Count	304	★	
	Sentence Count	19	★	
	Paragraph Count	6	★	
Language Issues				
😊	Spelling Issues	0 0%	★	