

# CSC 340: Programming Languages & Compilation

## Assignment\_1

### Regular Expressions

#### What is a web crawler used for?

An Email Extractor tool is a browser extension or web application that automatically extracts email addresses from websites, web pages, and social media platforms.

Web crawlers help you index web pages, locate content, and gather data from public websites. Crawlers also look at URL links within a website and try to work out a schema for how these pages are interconnected. This crawling helps you analyze the website from a wider perspective and makes it easier for online tools such as search engines to display a simplified version of it in the form of search results.

In this assignment you are request to use regular expressions based on web crawler platform to extract KSU e-mail addresses from the URL: <https://ccis.ksu.edu.sa/en> into a modified copy of a text file.

- 1) This assignment can be group of 2 students
- 2) Submit your pdf file and java file in one zip file on LMS (Make sure only ONE student of each group submits)
- 3) Your pdf file should contain:
  - i. A cover page (your name, serial number and section number)
  - ii. Java code and screen shots of your code and outputs
- 4) Due date (Sunday-18 February 2024 /23:59 PM)