****

**King Saud University**

**College of Computer and Information Sciences**

**Computer Science Department**

*CSC 340 – Assignment (1) report*

**Prepared by:**

(**8**) | Faisal Khalid Al-Zahrani | 441101794

**Supervised by:**

Dr. Wesam A. Hatamleh

2nd semester 1444

2022

**Section: 53505**

**Picture of the java code**

**Text

Description automatically generated**

**Java Code**

**package webCrawler;**

**import java.util.regex.Matcher;**

**import java.util.regex.Pattern;**

**import java.util.ArrayList;**

**import java.util.List;**

**import java.io.BufferedReader;**

**import java.io.InputStreamReader;**

**import java.net.URL;**

**public class webCrawler {**

**private List<String> emails1;**

**public static void main(String[] args) {**

**webCrawler wc = new webCrawler();**

**wc.crawl("https://ccis.ksu.edu.sa/ar/cs/faculty-and-staff");**

**}**

**public webCrawler() {**

**emails1 = new ArrayList<>();**

**}**

**public void crawl(String URL) {**

**String Hurl = "";**

**try {**

**URL url = new URL(URL);**

**BufferedReader n = new BufferedReader(new InputStreamReader(url.openStream()));**

**String inputLine = n.readLine();**

**while (inputLine != null) {**

**Hurl += inputLine;**

**inputLine = n.readLine();**

**}**

**n.close();**

**} catch (Exception e) {**

**e.printStackTrace();**

**}**

**String emailPattern = "(\\w+)(@ksu.edu.sa|@ccis.edu.sa)";**

**Pattern P = Pattern.*compile*(emailPattern);**

**Matcher M = P.matcher(Hurl);**

**while (M.find()) {**

**String email = M.group();**

**if (!emails1.contains(email))**

**emails1.add(email);**

**}**

**for (int i = 0; i < emails1.size(); i++)**

**System.*out*.println(emails1.get(i));**

**}**

**}**

**Picture of the output, THERE ARE 3 PICTURES**

**Text

Description automatically generated with medium confidenceA picture containing text

Description automatically generatedText

Description automatically generated**