**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**JATIYA KABI KAZI NAZRUL ISLAM UNIVERSITY, TRISHAL**

**Date**: 03-10-2019

1. **Name of the students**: Deena Faria Mahmuda Akhter

**Roll No**.: 17102005 17102027

**Session:** 2016-17 2016-17

1. **Programme:**  □PhD/□MS/□BSc.

**Proposal for:** □Thesis/□Project

1. **Name of the Proposed Supervisor:** Dr. Tushar Kanti Saha

**Designation**: □PROFESSOR / □ASSOCIATE PROFESSOR/ □ASSISTANT PROFESSOR/ □LECTURER

1. **Tentative Title:** JKKNIU News Grabber
2. **Background and present state of the problem:**

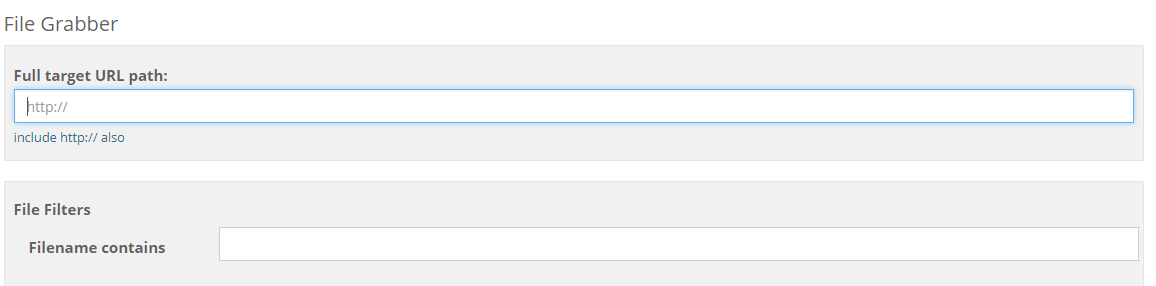
There are many different web sites or online news portal which can provide news about various things. We often require specific news from those. But for this reason, we have to search the web portals which can satisfy our request. This cost lots of time and data. But if we can make a news grabber, it can extract all the specific news we required from corresponding web pages. News grabber can grab, retrieve or download various media files from a public web site use them locally on client side.

1. **Objectives with specific aims and possible outcome:**

**Objective:**

1. Grab anything embedded or referred by the targeted url page depending on the content type HTTP response headers.
2. HTTP headers Content-Type filtering: only those files having the specific content-type will be included in the grabbing results.
3. Grab any document or file embedded or referred by the targeted url page depending on its type.
4. File name filtering: only those files containing the specified strings or extensions will be included in the grabbing results.

**Outcomes:**

****

The outcomes of this project will have two fields:

1. The URL of the desired web site
2. File name / Articles heading / Any string in the document
3. **Outline of Methodology/ Experimental Design:**
4. Create a news grabber which extract required document.
5. Create a user interface for this.
6. Requirements:
7. Apache web server.
8. Language: PHP, HTML, JavaScript etc.
9. **References:**
10. <http://wiseloop.com/tutorial/php-web-media-grabber/tutorial>
11. http://wiseloop.com/demo/php-web-media-grabber/6.demo-file-grabber