

(12) PATENT APPLICATION PUBLICATION

(21) Application No.202311045536 A

(19) INDIA

(22) Date of filing of Application :06/07/2023

(43) Publication Date : 04/08/2023

(54) Title of the invention : SYSTEM AND METHOD FOR ARCHIVING WEBSITES ON A LOCAL MACHINE

(51) International classification :G06F 030484, G06F 084100, G06F 169535, G11B 270340, H04L 093200
(86) International Application No :NA
Filing Date :NA
(87) International Publication No : NA
(61) Patent of Addition to Application Number :NA
Filing Date :NA
(62) Divisional to Application Number :NA
Filing Date :NA

(71)Name of Applicant :

1)Chitkara University

Address of Applicant :Chitkara University, Chandigarh-Patiala National Highway, Village Jhansla, Rajpura, Punjab - 140401, India. Patiala -----

2)Bluest Mettle Solutions Private Limited

Name of Applicant : NA

Address of Applicant : NA

(72)Name of Inventor :

1)MISHRA, Saket

Address of Applicant :ODC-4, Panchshil Tech Park, inside Courtyard by Marriott premises, Hinjewadi Phase - 1, Pune - 411057, Maharashtra, India. Pune -----

2)PANDEY, Sakshi

Address of Applicant :ODC-4, Panchshil Tech Park, inside Courtyard by Marriott premises, Hinjewadi Phase - 1, Pune - 411057, Maharashtra, India. Pune -----

3)GILL, Rupali

Address of Applicant :Chitkara University, Chandigarh-Patiala National Highway, Village Jhansla, Rajpura, Punjab - 140401, India. Patiala -----

(57) Abstract :

The present invention discloses a system (100) and a method (200) for archiving a plurality of websites on a local machine. The system includes a processor (102) that includes an archiving module that extracts a plurality of materials from the websites and stores them in the memory, and a compression module compresses the extracted web pages to optimize storage space. A user input a set of URLs and configures archiving options using a computing device, and the system enables automatic import and crawling of the URLs at predefined intervals using a scheduling and crawling module. Multiple snapshots of the URLs are generated in formats such as HTML, PDF, PNG screenshots, and WARC. The system also facilitates the sharing of archived web pages with one or more entities.

No. of Pages : 25 No. of Claims : 10