Názov výsledku: Systém pre zber a monitorovanie dát z dynamických webových stránok
Anglicky názov výsledku: System for collecting and monitoring data from dynamic web pages
Autori: Emil Gatial

Anotácia výsledku

Hlavná motivácia pre návrh a tvorbu zberného a monitorovacieho systému je umožniť bežnému používateľovi sledovať zmeny bez nutnosti použiť komplikované alebo drahé riešenia. Hlavnou výhodou systému je použitie zásuvného modulu (plugin) do bežného internetového prehliadača (Firefox alebo Chorme) pre anotovanie a zadávanie parametrov získavania, ktorý zefektívňuje zadávanie parametrov bez nutnosti prepínať sa do externej aplikácie. Systém pozostáva z dvoch softvérových modulov: zásuvný modul Internetového prehliadača poskytujúci používateľské grafické rozhranie (GUI) a modul pre extrakčné a notifikačné služby. Prvý modul umožňuje používateľovi vytvoriť „predpis“ pre extrakciu dátového elementu, ktorý pozostáva z umiestnenia elementu (Xpath, URL adresa), plánu extrakcie (štart, ukončenie a frekvencie extrakcie) a notifikačné parametre (treshold hodnoty, notifikačný kanál). Druhý modul poskytuje prístup k aplikačnej logike systému prostredníctvom REST služieb a asynchrónneho komunikačného kanála Web Socket. Tento modul pozostáva z plánovača úloh, paralelne bežiacich extrakčných procesov a škálovateľného úložiska (Appache Cassandra). Pre testovanie boli vytvorené dva testovacie scenáre. Prvý scenár má za úlohu sledovať vývoj cien na akciovom trhu. Druhý scenár má za úlohu monitorovať ceny leteniek viacerých leteckých spoločností prostredníctvom sprostredkovateľských portálov. Systém je využiteľný aj na testovanie a monitorovanie webových stránok.

Hlavné scientometrické výstupy

1. Gatial E., Balogh Z., Hluchý L.: Change detection and notification method of the rich Internet application content. In: Proceedings of INES 2018 (IEEE 22nd International Conference on Intelligent Engineering Systems). – Óbuda University, Budapest, Hungary, 2018, p. 51-56. ISBN: 978-1-5386-1121-0