

Title (en)

WEB SCRAPING THROUGH USE OF PROXIES, AND APPLICATIONS THEREOF

Title (de)

BAHNABKRATZEN DURCH VERWENDUNG VON PROXIES UND ANWENDUNGEN DAVON

Title (fr)

MOISSONNAGE DU WEB AU MOYEN DE SERVEURS MANDATAIRES, ET APPLICATIONS CORRESPONDANTES

Publication

EP 4222617 A1 20230809 (EN)

Application

EP 22740786 A 20220624

Priority

- US 202163219660 P 20210708
- US 202117373287 A 20210712
- US 202117373312 A 20210712
- US 202117373482 A 20210712
- US 202117373570 A 20210712
- US 202117373608 A 20210712
- US 202117373634 A 20210712
- EP 2022067331 W 20220624

Abstract (en)

[origin: EP4227828A1] Embodiments disclose a system that allows for improved generation of web requests for scraping that, because of the nature of the requests and time and manner they are sent out, appear more organic, as in human generated, than conventional automated scraping systems. The system then manages how a client request to scrape a target website is made to the site, masking the request in a manner that makes it appear to the Web server as if the request is not generated by an automated system. In this way, by appearing more organic, Web servers may be less likely to block requests from the disclosed system or may take longer to block requests from the disclosed system. By avoiding Web servers blocking requests and extending the lifetime of IP proxies before they are blocked, embodiments can use a limited IP proxy address space more efficiently.

IPC 8 full level

G06F 16/958 (2019.01); **G06F 16/951** (2019.01)

CPC (source: EP IL)

G06F 16/951 (2019.01 - EP IL); **G06F 16/972** (2019.01 - EP IL)

Designated contracting state (EPC)

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Designated extension state (EPC)

BA ME

Designated validation state (EPC)

KH MA MD TN

DOCDB simple family (publication)

EP 4227828 A1 20230816; CA 3214799 A1 20230112; EP 4222617 A1 20230809; EP 4227829 A1 20230816; IL 308559 A 20240101; MX 2023015142 A 20240122

DOCDB simple family (application)

EP 23171587 A 20220624; CA 3214799 A 20220624; EP 22740786 A 20220624; EP 23171700 A 20220624; IL 30855923 A 20231114; MX 2023015142 A 20220624