

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
12 January 2023 (12.01.2023)



(10) International Publication Number
WO 2023/280593 A1

(51) International Patent Classification:
G06F 16/958 (2019.01) *G06F 16/951* (2019.01)

(21) International Application Number:
PCT/EP2022/067331

(22) International Filing Date:
24 June 2022 (24.06.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/219,660	08 July 2021 (08.07.2021)	US
17/373,287	12 July 2021 (12.07.2021)	US
17/373,312	12 July 2021 (12.07.2021)	US
17/373,482	12 July 2021 (12.07.2021)	US
17/373,570	12 July 2021 (12.07.2021)	US
17/373,608	12 July 2021 (12.07.2021)	US
17/373,634	12 July 2021 (12.07.2021)	US

(71) Applicant: **METACLUSTER LT, UAB** [LT/LT]; A.
Gostauto g. 40A, 03163 Vilnius (LT).

(72) Inventors: **VILCINSKAS, Eivyd**as; Korsako Street 15-6,
78358 Siauliai City (LT). **PETRUŠKEVICIUS, Arn**as;
Kalvariju g. 137D-27, 08221 Vilnius (LT). **STALIO-
RAITIS, Giedri**us; Burbiškių 60-14, 02198 Vilnius (LT).

JURAVICIUS, Martynas; Naugarduko Street 62-85,
03203 Vilnius (LT). **STANKEVICIUS, Rimantas**; Bi-
tininku street 4F-23, 08310 Vilnius (LT).

(74) Agent: **MEISSNER BOLTE PATENTANWÄLTE
RECHTSANWÄLTE PARTNERSCHAFT MBB**;
MEYER, Florian, Widenmayerstr. 47, 80538 Munich (DE).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA,
MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,
RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM,
ZW.

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,

(54) Title: WEB SCRAPING THROUGH USE OF PROXIES, AND APPLICATIONS THEREOF

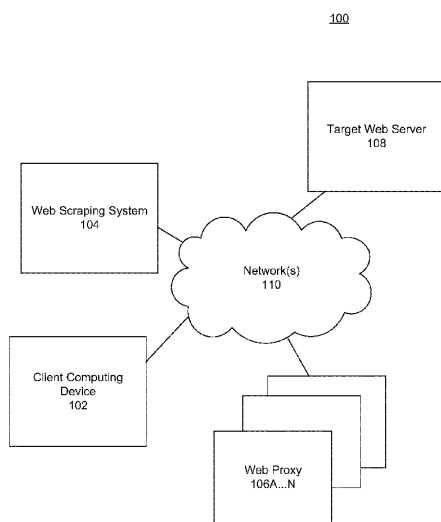


FIG. 1

(57) Abstract: 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.

WO 2023/280593 A1

[Continued on next page]

MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *with amended claims (Art. 19(1))*
- *in black and white; the international application as filed
contained color or greyscale and is available for download
from PATENTSCOPE*