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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 November 2013 (07.11.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/163964 A1

(51) International Patent Classification:
G06F 17/28 (2006.01)

(21) International Application Number:
PCT/CY2012/000002

(22) International Filing Date:
2 May 2012 (02.05.2012)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **XCRE-MENO LTD** [CY/CY]; 35, Thekla Lysiotis House, 6th Floor, CY-3030 Limassol (CY).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **ANDERSSON, Daniel** [SE/CY]; 35, Thekla Lysiotis Str., Eagle Star House, 6th Floor, CY-3030 Limassol (CY).

(74) Agent: **MICHAEL KYPRIANOU & CO LLC**; 2 Agias Elenis Str., Stasinos Building, 6th & 7th Floor, CY-1060 Nicosia (CY).

(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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, 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, 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, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: WEBSITE TRANSLATION DELIVERY AND MANIPULATION

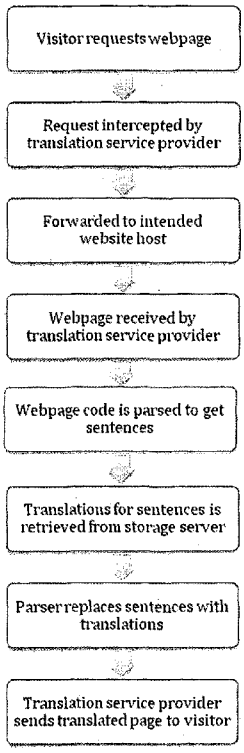

Fig. 1

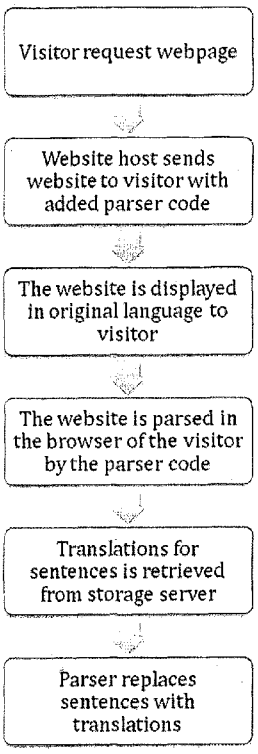

Fig. 2

(57) Abstract: A method of delivering and manipulating content on a web page in near real time by a third party service that is compatible with any web page to deliver a translation. The web page may be displayed in the original language initially until the translation is available.

WO 2013/163964 A1

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

WEBSITE TRANSLATION DELIVERY AND MANIPULATION

Background of the invention

[0001] 1. Field of the Invention

[00] The present invention relates to website translation delivery. More particularly it relates to how translations are stored, accessed and delivered from a

third party while a user is accessing a web page online.

[00] 2. Description of Related Art

[00] Today there exist many types of translation systems to store translation. These are custom systems tied into generally pre-defined content management system that is operating on the host computer of a website. These content management systems generally has pre-defined capabilities such as news content, articles, shopping items tied in with text. The translation system is then applied to each capability and is operating to translate these items.

[00] Translation systems for web pages today allows for a website to be available in multiple languages by having multiple copies of content available tied together in a backend system. By example, a site might have a database full of news articles, instead of having it stored as an index such as News ID & Text, it's stored as a News ID & English Text, German Text. There's then a separate function that stores other parts of the website, such as menu, about, footer information, banners, etc in the same manner. When a website is served it calls up on these different custom functions to assemble a fully functional website in the language a user requested.

[00] By design current translation systems are always tied in to content, that will say to create a new function, such as a website evolving to create a news feed for it's visitors, it's then always required to also create a custom function that can tie this news content into a translation system. Increasing the complexity of the backend system powering a website exponentially.

[00] The problem found in current systems, where the translation is tied together with each part of a website, is the lack of scalability and the increased complexity to produce a new function for a website in a new language. Because there's no standard for translations or content management there can't be a system that fits the backend of every webpage due the custom nature of website functions and features.

[00] The only current standardized webpage translation tools comes part of a package deal for website backend systems, where you have a preset amount of features such as new and articles where the translation system will support these functions and no other, unless extended and customized by the end user creating new system functionality, if allowed by the license of the system used.

Summary of the invention

[00] The invention relates to an external translation mechanism that will be positioned as a third party service and host translations on a separate host. Whenever a visitor requests a webpage there will be a second request to an external server that will retrieve translations for a webpage. There will also be a system in-between the visitor and the original host or a system executed locally on the visitor's browser that will manipulate the webpage code to insert any translated text in place of the original text. The end result will be a full translation for any webpage independent of the underlying system or programming language used to generate the web page HTML code by the original host.

[00] Because this system is offered as a third party service, and is executed after the fact of the website host and independent of the website host programming language and system functionalities it will work with any website. It is capable of extracting the text out of the web page the visitor

requests, find any translation for said text, and then replace the text in the web page code, and forward it to the user in near real time.

Brief description of the drawings

[00]

Detailed description of the invention

[00] The invented method of website translation stores every sentence and its designated translation in a third party system that will offer near real time translations as a service. The third party system could simultaneously keep translations for millions of web pages of thousands of different unique websites and offer each sentence up as requested.

[00] By example, there are currently two different methods of delivering these translations to a visitor. The first method , illustrated in **Fig. 1** places an http proxy in-between the visitor and the website host. The website owner will have to arrange for two things in order to setup the http proxy for translation. A specific domain name, such as de.example.com for German would be forwarded via it's DNS settings to the translation service providers IP, and the owner would then specify where the source website is. Which in this example would be an English website located on www.example.com. In effect, what happens when a visitor visits de.example.com/example-path1.html, is that the visitor requests a webpage from the translation service provider, the system would in turn contact the source at www.example.com/example-path1.html, download the full web page, parse the code to split and find every translation. It would then put the webpage together again with the German text replacing the English text, and then forwarded to the visitor just as if the visitor requested the web page from the original host, but fully translated.

[00] The second method, illustrated in **Fig. 2** involves an external script being included ißn the webpage code. When a webpage is loaded, it will be sent to the visitor from the original host, and include an extra line of external code that will be executed in the browser of the visitor. When the webpage is requested, sent and displayed to the visitor in its original language, this code will be executed on the visitor's browser to parse and find all text elements, which will be sent to the third party service provider, which will find all translations for these sentences, and send it back. The software code would then switch out text on the webpage with the translation. What this second method provides is an easier integration for the website owner, in return the webpage will be displayed briefly, usually for less than 1 or 2 seconds, to the user in its original language and then the text will be switched out with the translated counterpart in front of the eyes of the visitor on every web page.

[00] The background parsing element for a website works in essentially the same way for both the first and second method. The HTML code of a website, which is the standard of any website, can be broken down to separate each text element from any other part of the HTML code of a webpage. Once this is done there's a list of sentences that is tied back to its position in the page. These sentences as an array is sent to a translation delivery system combined with the unique location on the web – by example de.example.com/example-path.html. Once the translation storage server receives both these data points, it can retrieve the translations for this path by the sentences it received and reply with a new array that

contains. When the parser receives this translation array, which maintains the pointers to where in the original HTML code each sentence is located, can then go thru every sentence, remove the original sentence and put in the translated sentence in its place. The result would be the exact same web page but with every sentence translated.

[00] The storage server of translations would maintain an index, which every single webpage that would have a sub index of every sentence of that webpage. For immediate access, this full database will be stored in RAM memory of the host server, which will maintain a retrieval time of milliseconds per sentence in the range of 5-50 milliseconds for the sentences of a full webpage when accessed in parallel, to successfully match the speed expectations by a web page visitor. When requested, it can match the sentences sent with the sentences stored and if every sentences requested exists as a translation it can send and complete the request. In case the storage server finds that there are new sentences sent by the parser which is not stored, there are three different outcomes that can be pre-defined by the website owner on a per-website scope. What would be constant every time a sentence is missed is that it will be logged and alerted to a translation index server which can then in turn alert and request a translation for this sentence, in effect automating detection of any changes on a web page that can be automatically sent out to be translated.

[00] The first option would send the full website to the visitor which would translate as many words as available, and then leave the words not yet translated.

[00] The second option would create an index of every sentence not yet translated and request an automatic translation from any available service, such as Google Translate. Once the automatic translation is completed it can be mixed together with the original translation and the request can be completed and sent to the parser and in turn to the visitor. After this has been sent the automatic translated text can be put in to the translation storage database with the explicit tagged as automatic, allowing them to be subsequently served for any future requests while also being tagged as temporary and scheduled to be translated by a real human translator.

[00] The third option can display an information box for the user that the current page is not yet translated in full, who can then choose to view the webpage in it's original language, or as the first option, or as the second option.

Claims

1. This invention is a unique way to provide for website translation delivery and manipulation.
2. It provides a solution for the lack of scalability found in current website translation methods.
3. Tied to the claim number 1 is that this invention is related to an external translation mechanism that will be positioned as a third party service and host translations on a separate host.
4. There will be a connection between the visitor and the original host that will manipulate the website code to insert any translated text in place of the original text.
5. This method is unique and independent. This is because it is offered as a third party service and will be executed after the website host and it will be independent from the website host programming language as well as any system functionalities. This means that it will work with any website.
6. In addition another way is by including an external script in the webpage code. When webpage is loaded, it will be sent to the visitor from the original host, and include an extra line of external code that will be executed in the browser of the visitor
7. It will extract the text out of the web page the visitor requests, find a translation for the said text and then replace the text in the webpage code, and forward it to the user in near real time.
8. The storage server of translations would maintain an index, with every single webpage and in turn a sub index if every sentence of that webpage. For immediate access, this full database will be stored in RAM memory of the host server, which will have a retrieval time of milliseconds per sentence.

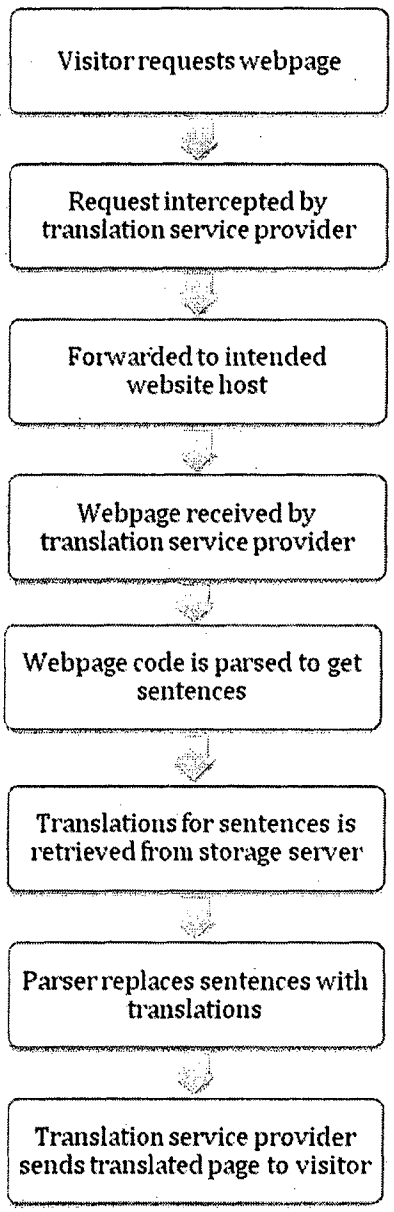
Visitor requests webpage
Request intercepted by translation service provider
Forwarded to intended website host
Webpage received by translation service provider
Webpage code is parsed to get sentences
Translations for sentences is retrieved from storage server
Parser replaces sentences with translations
Translation service provider sends translated page to visitor

Fig. 1

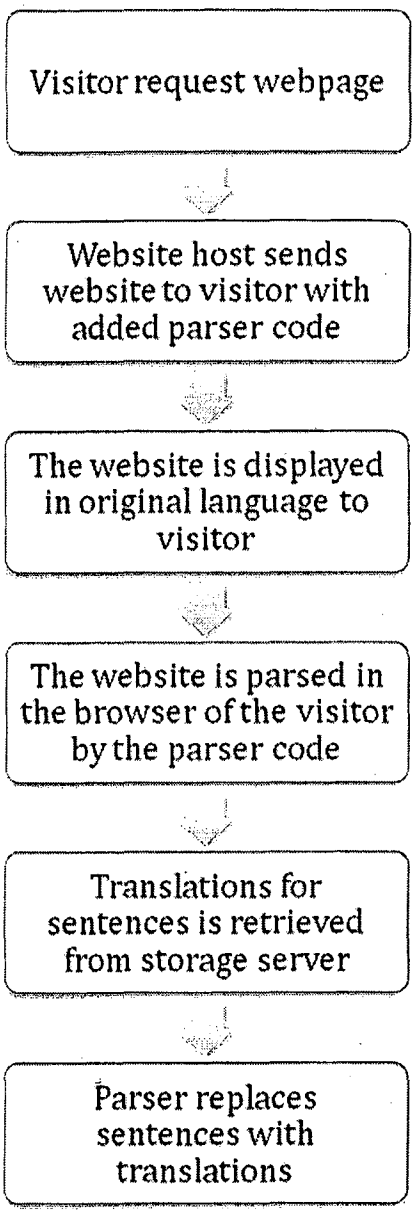
Visitor request webpage
Website host sends website to visitor with added parser code
The website is displayed in original language to visitor
The website is parsed in the browser of the visitor by the parser code
Translations for sentences is retrieved from storage server
Parser replaces sentences with translations

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/CY2012/000002

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/28
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2012/016655 A1 (TRAVIESO ENRIQUE [US] ET AL) 19 January 2012 (2012-01-19) abstract paragraphs [0052] - [0073]; figures 5, 6a-c, 7a	1,3-5,7, 8 6
Y	US 2008/077384 A1 (AGAPI CIPRIAN [US] ET AL) 27 March 2008 (2008-03-27) abstract paragraphs [0001] - [0015]	6

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 February 2013	14/02/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Woods, Justin

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CY2012/000002

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 2(completely); 1, 3-8(partially)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 2(completely); 1, 3-8(partially)

Contrary to Article 6 and Rule 6.3 PCT, by merely referring to the technical problem, without setting out a solution to same, claim 1 fails to define the matter for which protection is sought in a manner which is clear and concise and in terms of the technical features of the invention. The subject matter of claim 1 is so thus broad, that no meaningful search into the prior art can be carried out in its respect because such a search would have to include all instances of such a problem, regardless of the solution applied. However, the description points to two methods of delivering a translation, respectively set out in figures 1 and 2 and broadly corresponding to claims 1 and 6 respectively. Thus claim 1 is interpreted in the light of the description, with particular reference to the features set out in the flow chart of claim 1. Claim 6 is interpreted in the light of the description as an alternative method of website delivery described in the description and is interpreted using the steps shown in figure 2 (which has been erroneously included as part of the description). Accordingly, an incomplete search of claim 1 has been made using this interpretation. This approach enables an interpretation of claims 3-5, 7 and 8 as being dependent on claim 1 even thought they have not been explicitly designated as dependent claims. Accordingly, an incomplete search of claims 3-5, 7 and 8 has been made using this interpretation. Claim 2 merely relates to a result to be achieved and cannot be searched at all. Even if interpreted as being dependent on claim 1, claim 2 merely sets out a result to be achieved, without providing any indication of how the effect is to be achieved, thus rendering the claim unclear in the sense of Article 6 PCT. As the objection relates to the entirety of the claim, no meaningful search could be carried out w.r.t. to its subject matter.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.2), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/CY2012/000002

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012016655 A1	19-01-2012	US 2012016655 A1	19-01-2012
		US 2012016656 A1	19-01-2012
		US 2012016865 A1	19-01-2012
		US 2012016929 A1	19-01-2012
		US 2012017146 A1	19-01-2012
		WO 2012009441 A2	19-01-2012
US 2008077384 A1	27-03-2008	NONE	