



A Micro Crowdsourcing Architecture to Localize Web Content for Less-Resourced Languages

Asanka Wasala

Chris Exton

Ruvan Weerasinghe

Reinhard Schäler

English

French
German
Italian
Spanish
Castilian
Portuguese
Turkish
Polish
Russian
Chinese
Japanese
Arabic
Flemish
Persian
Dutch
Swedish
Norwegian

Adapted from: http://w3techs.com/technologies/overview/content_language/all



Adapted from: http://w3techs.com/technologies/overview/content_language/all

Scenario

- You've come across a Sinhala web site..



විකිපීඩියා
නිදහස් විශ්වකෝෂය

- මුල් පිටුව
- ප්‍රජා ද්වාරය
- කාලිනා සිදුවීම්
- නව වෙනස්කිරීම්
- අහඹු පිටුව
- උදව්
- පරිත්‍යාග
- මෙවලම් ගොන්න
- සබැඳි පිටු
- සහසම්බන්ධිත වෙනස්වීම්
- විශේෂ පිටු
- මුද්‍රණ ආකෘතිය
- ස්ථාවර සබැඳුම්
- මෙම පිටුව උපන්‍යාස කරන්න

වෙනත් භාෂා වලින්
English

පිටුව සාකච්ඡාව කියවන්න සංස්කරණය කරන්න ඉතිහාසය නරඹන්න ගවේෂණය කරන්න

If you have trouble reading sinhala fonts, [click here](#).

සිංහල හින්දු අලුත් අවුරුද්ද

විකිපීඩියා, නිදහස් විශ්වකෝෂය වෙතින්

සිංහල අලුත් අවුරුද්ද සහ දෙමළ අලුත් අවුරුද්ද වෙනස් වාරික ඉටුකරණ උත්සව වන අතර , නව වසර , **අබංගාලි නව වසර** වැනි දකුණු සහ අග්නිදිග ආසියානු අලුත් අවුරුද්ද සමරයි.

මෙයට අදාල නිවාඩු දිනය ශ්‍රී ලංකා දින දසුනේ සටහන් කරන්නේ **සිංහල සහ හින්දු අලුත් අවුරුද්ද** නමිනි.

මෙම නව වසර සංස්කෘතික වශයෙන් අප්‍රේල් හෙවත් ඔක් මාසයේ මාසයේ සමරනු ලබන අතර අතීතයේ සිටම ශ්‍රී ලංකාවේ පැවති මහා කලා මංගල්‍යයකි.විවිධ විශ්වාස සිංහල හින්දු අලුත් අවුරුද්ද හා බැඳී පවතී.කෘෂිකාර්මික කටයුතු හා බැඳුණු විශ්වාස, වාරිත්‍ර වාරිත්‍ර, ආගමික විශ්වාස මේ අතර වේ. බුද්ධ වර්ෂ 300 දී පමණ ශ්‍රී ලංකාවට බුදු දහම හඳුන්වාදීම මඟින් බෞද්ධාගමික පරිසරයක් තුල අලුත් අවුරුදු වාරිත්‍ර කිරීමට මහ පෙත්වනු ලැබීය.රටේ බහුතර ජනතාව බෞද්ධයන් වන නිසා බෞද්ධ දර්ශනයට ප්‍රමුඛත්වය දෙමින් සිංහල අවුරුදු වත් පිළිවෙත් කිරීමට අදත් ශ්‍රී ලාංකික ජනතාව හුරු පුරුදු වී සිටිති.

අනෙක් අතට හින්දු ආගම ගැන සලකා බලන කල එය පුරාතණයේ සිටම බුදු දහම හා සමාන්තරව පැවතුනු ආගමක් ලෙස හඳුනා ගත හැක.එම නිසා නව අවුරුදු වාරිත්‍ර වාරිත්‍ර හින්දු ජනතාව අතර හින්දු ආගමට අනුකූලව වැඩි දියුණු විය.බෞද්ධාගම හා හින්දු ආගම ඓතිහාසිකව එකිනෙක හා බැඳී පැවතිනි. එම නිසා බෞද්ධයින් හා හින්දු භක්තිකයන් අතර අවුරුදු වාරිත්‍ර වල විශාල පරස්පරතා හඳුනා ගත නොහැක.

මීර්‍යා විශ්වාස හා ප්‍රබන්ධ වලට අනුව අලුත් අවුරුදු උදාව හින්දු සාහිත්‍යයට නියත වශයෙන්ම බැඳී පවතී.සාමයට අධිපති කුමරා ලෙස හින්දු ආගමේ සලකනු ලබන 'ඉන්ද්‍රදේව'පොළව වෙත සාමය හා සෞභාග්‍යය ලබා දීම තහවුරු කිරීම ලෙස නව අවුරුදු උදාව සැලකේ.අවුරුදු කුමරා සුදු පැහැ රථයක නැඟී සුදු මල් වලින් සරසනු ලැබූ හත් රියන් උස ඔවුන්නක් හිස පැලඳ පොළවට වඩයි.ඔහු පළමුව පොළවේ ගුරුත්වාකර්ෂණ බලය බිඳ දමා එය කිරි සුදුරු හෙවත් කිරි මුහුදේ ගිල්වා ගොඩට ගනු ලබන බවට මිත්‍යා විශ්වාසයක් ගැමි ජනයා අතර පවතී.



සිහි අලුත් අවුරුද්දක් වේවා
සාමය සතුට සම්පත පිරි
වාසනාවන්ත තව



File Edit View History Bookmarks Tools Help



www.translate.google.com

Most Visited Getting Started Latest Headlines

Google Translate

Web Images Videos Maps News Shopping Gmail more ▾

Sign in ⚙



Translate

From: Detect language ▾



To: English ▾

Translate

http://si.wikipedia.org/w

Type text or a website address or file

Detect language	Chinese	Georgian	Italian	Persian	Tamil
Afrikaans	Croatian	German	Japanese	Polish	Telugu
Albanian	Czech	Greek	Kannada	Portuguese	Thai
Arabic	Danish	Gujarati	Korean	Romanian	Turkish
Armenian	Dutch	Haitian Creole	Latin	Russian	Ukrainian
Azerbaijani	English	Hebrew	Latvian	Serbian	Urdu
Basque	Estonian	Hindi	Lithuanian	Slovak	Vietnamese
Belarusian	Filipino	Hungarian	Macedonian	Slovenian	Welsh
Bengali	Finnish	Icelandic	Malay	Spanish	Yiddish
Bulgarian	French	Indonesian	Maltese	Swahili	
Catalan	Galician	Irish	Norwegian	Swedish	

Translations: [Dismiss](#)

Google Translate for my: [Searches](#) [Videos](#) [Email](#) [Phone](#) [Chat](#) Business: [Translator Toolkit](#) [Global Market Finder](#) [Website Translator](#)

[About Google Translate](#) [Turn off instant translation](#) [Privacy](#) [Help](#)

Bing Translator

Web Images News More | MSN Hotmail

Sign in Preferences

bing
Translator

Home | Users | Developers | Webmasters | Help

Translate from: Auto-Detect

Translate

Translate to:

- Auto-Detect
- Indonesian
- Italian
- Arabic
- Japanese
- Bulgarian
- Korean
- Catalan
- Latvian
- Chinese Simplified
- Lithuanian
- Chinese Traditional
- Norwegian
- Czech
- Polish
- Danish
- Portuguese
- Dutch
- Romanian
- English
- Russian
- Estonian
- Slovak
- Finnish
- Slovenian
- French
- Spanish
- German
- Swedish
- Greek
- Thai
- Haitian Creole
- Turkish
- Hebrew
- Ukrainian
- Hindi
- Vietnamese
- Hungarian

http://si.w

ဘီ-အိတ်-အိတ်

Enter text or w

Language Tool

See this page o

Microsoft® Indic

Bing Dictionary

Microsoft® Mar

Powered by Microsoft® Translator



The Microsoft® Translator [webpage widget](#) allows you to bring real-time, in-place translations to your web site. The collaborative translation features combine automatic translation technology with the power of your community. Try the all new translator invitation and translation management tools!



Tbot is an automated buddy that provides translations for Windows Live Messenger. You can have one-on-one conversations with Tbot or invite friends who speak different languages with Tbot translating for you. Simply add mtbot@hotmail.com to your Messenger contacts.

© 2011 Microsoft | Privacy | Legal | Advertise | Help | Feedback



New User? Register | Sign In | Help

Yahoo! Mail My

YAHOO! BABEL FISH

Search

Web Search

Translate a block of text (Enter up to 150 words)

Select from and to languages

Translate

Translate a web page

http://si.wikipedia.org/wiki/සිංහල_නින්ද_අලුත්_අවුරුද්ද

Select from and to languages

Translate

Select from and to languages

Chinese-simp to English
Chinese-simp to Chinese-trad
Chinese-trad to English
Chinese-trad to Chinese-simp
English to Chinese-simp
English to Chinese-trad
English to Dutch
English to French
English to German
English to Greek
English to Italian
English to Japanese
English to Korean
English to Portuguese
English to Russian
English to Spanish
Dutch to English

grammar, and punctuation for the highest

some text, click the button marked "Search the
order to launch a search using the translation

l web page with the original by clicking "View
language."



Web site Localisation..

Web site l10n

Automatic (MT)

Manual



Internet



English Web

Web site localisation..

“English dominated web pages on the Internet have resulted in businesses losing sales revenues of \$10 million per year”

-Harvard-based Forrester Research

English





WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[විශේෂාංගය content](#)
[Current events](#)
[Random article](#)
[තෑගි දෙනවා to Wikipedia](#)

▼ [Interaction](#)

[Help](#)
[About Wikipedia](#)
[Community ද්වාරය](#)
[Recent changes](#)
[Contact Wikipedia](#)

[Article](#)

[Discussion](#)

[Read](#)

[Edit](#)

[View history](#)



Sinhalese New Year

From Wikipedia, the free විශ්වකෝෂය



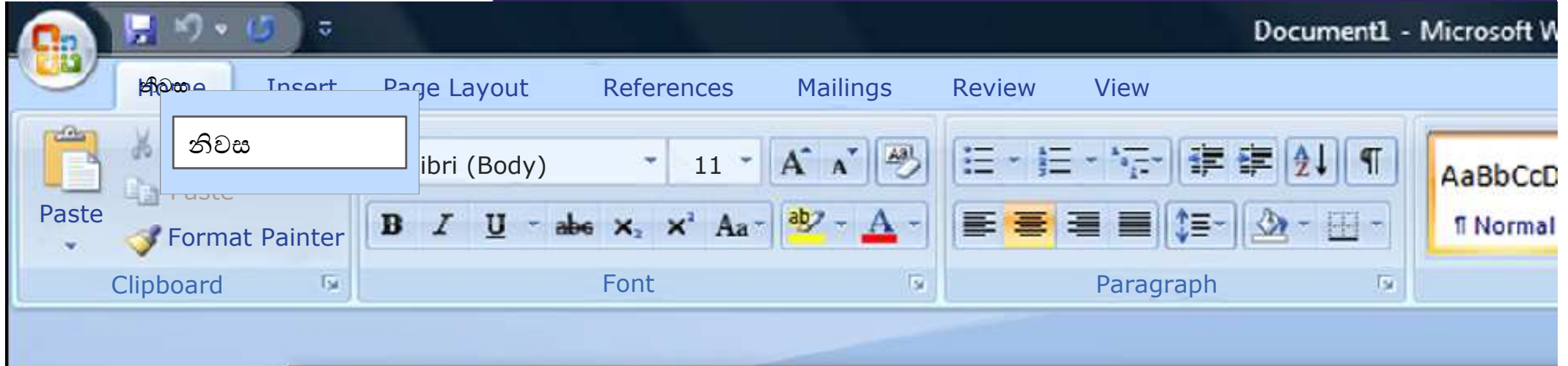
This article includes a [list of references](#), related reading or [external links](#), but **its ප්‍රභව remain unclear because it lacks inline උපහරණය**. Please [improve](#) this article by introducing more නියත citations [where උචිත](#). *(April 2011)*

The timing of the **Sinhala** New Year coincides with the new year උත්සවය of many පාරම්පරික calendars of **South** and **Southeast Asia**. The festival has close සාමාන්‍ය to the **Thai New year**, **Bengali New Year**, and **Sankranti** මංගල්‍යය in India^[*citation needed*].

සංස්කෘතික anthropological history of the 'පාරම්පරික New Year' which is celebrated on month of April, goes back to an ancient period in Sri Lankan history. Various ඇදහීම, perhaps those associated with සරුබව of the harvest, gave birth to many පූජාවිධි, customs, and ceremonies connected with the New Year. The පැමිණීම of **Buddhism** in the third

Sinhala new year

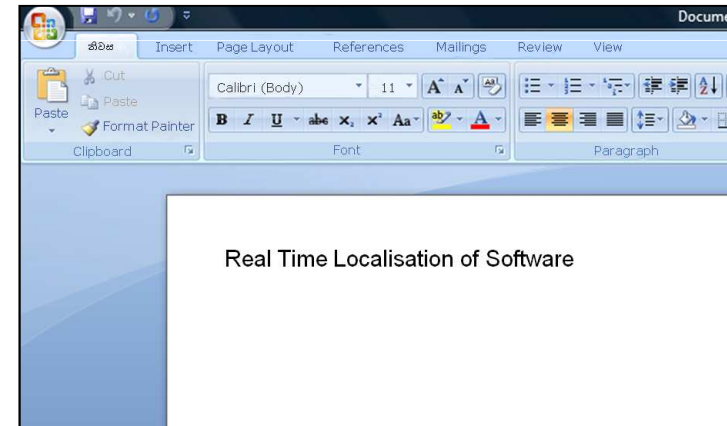
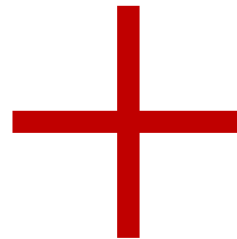
Official name	Sinhala Aluth Avurudda
Observed by	Sinhalese people
Type	Public holidays in Sri Lanka
Date	A nakshatra date in the month of Bak (by the Shalivahana era)
Celebrations	Games, prayer
Observances	The observed movement of the



Real Time Localisation of Software

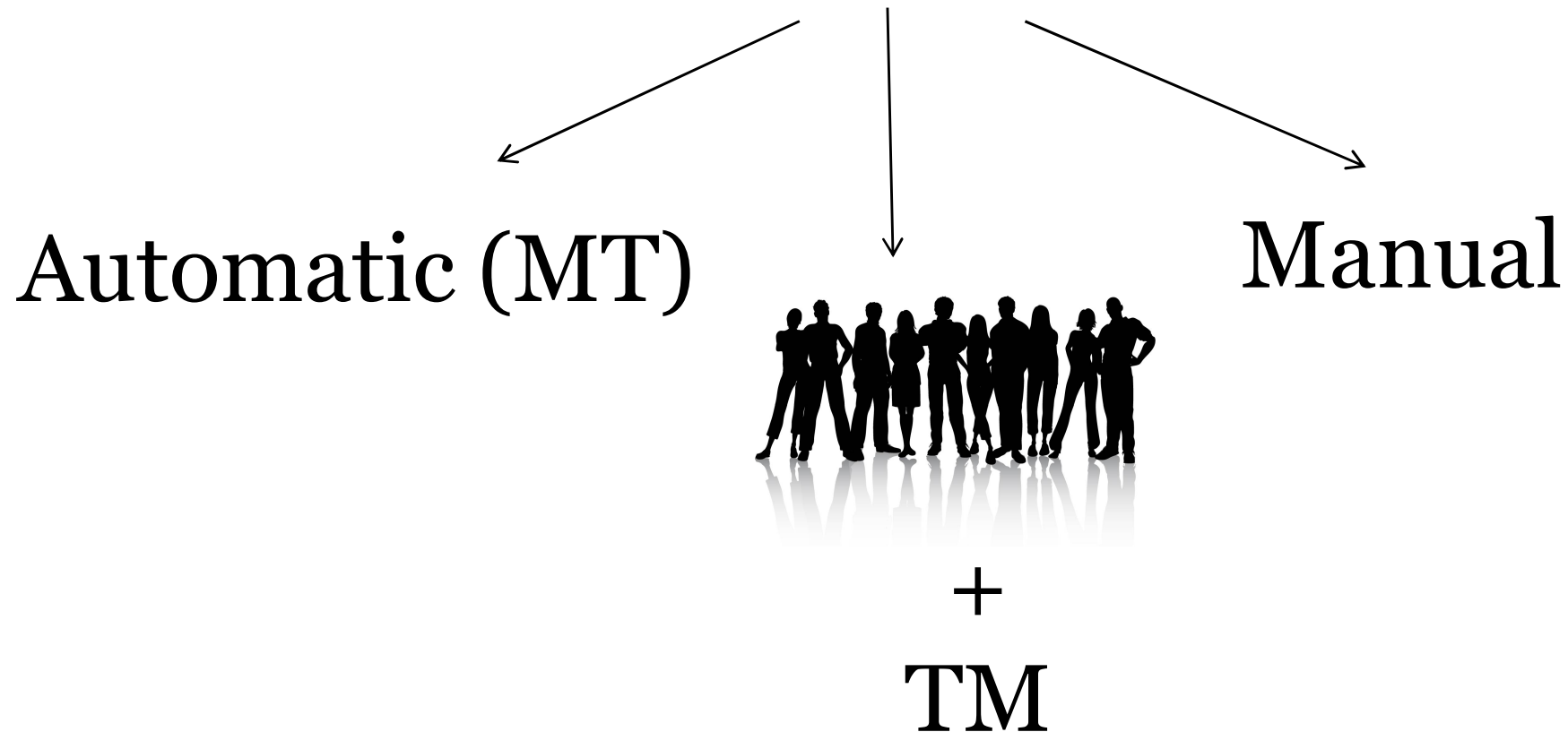
“Micro Crowdsourcing: A new Model for Software Localisation”. (2009).
Localisation Focus, The International Journal of Localisation ,Vol 8 Issue 1
Chris Exton, Asanka Wasala, Reinhard Schäler, Jim Buckley

Real Time Web Content Localisation



Web site Localisation..

Web site l10n





File Edit View History Bookmarks Tools Help



http://www.translationautomation.com/perspectives/lets-make-music-together.html



Most Visited Getting Started Latest Headlines

Let's Make Music Together!



HOME ABOUT TAUS PUBLICATIONS GUIDELINES EVENTS MEMBERS TRACKER STORE

Reports Translation in the 21st Century Best practices Technology Perspectives User cases

Home > Publications > Perspectives > Let's Make Music Together!

Let's Make Music Together!

FRIDAY, 08 APRIL 2011 09:52 | RUSTIN GIBBS

Facebook Share Like 14 Twitter Tweet 9 +1 0 LinkedIn Share 24 Submit

What is interoperability? Why is it important? How do we achieve



A few years ago, I converted my garage studio to support my musical hobbies. My initial investment I made in time and money that I have spent countless hours of experimenting with music on my own, my hobbies immensely, I always come to the same conclusion after each session—music is social; I need other musicians to help me fully experience the phenomenon of creative expression through music. I need a group. I need a band.

Open Link in New Tab
Open Link in New Window

Bookmark This Link
Save Link As...
Send Link...
Copy Link Location
Copy

Translate

JOIN OUR MAILING LIST

Enter email address

Submit

Find and shortlist machine solutions for your specific

TAUS host

TRANSLATION DEBATE



File Edit View History Bookmarks Tools Help

http://www.translationautomation.com/perspectives/lets-make-music-together.html

Most Visited Getting Started Latest Headlines

Let's Make Music Together!



HOME ABOUT TAUS PUBLICATIONS GUIDELINES EVENTS MEMBERS TRACKER STORE

Reports Translation in the 21st Century Best practices Technology Perspectives User cases

Home > Publications > Perspectives > Let's Make Music Together!

සැවොම එකතුවී ගී ගයමු

FRIDAY, 08 APRIL 2011 09:52 | RUSTIN GIBBS

Share Like 14 Tweet 9 +1 0 in Share 24 Submit

What is interoperability? Why is it important? How do we achieve it?



A few years ago, I converted my garage into a music studio to support my musical hobbies and interests. The initial investment I made in time and money has paid off in that I have spent countless hours of creative bliss experimenting with music on my own, but although I enjoy my hobbies immensely, I always come to the same conclusion after each session—music is social; I need other musicians to help me fully experience the phenomenon of creative expression through music. I need a group. I need a band.

JOIN OUR MAILING LIST

Enter email address

Submit

TAUS TRACKER

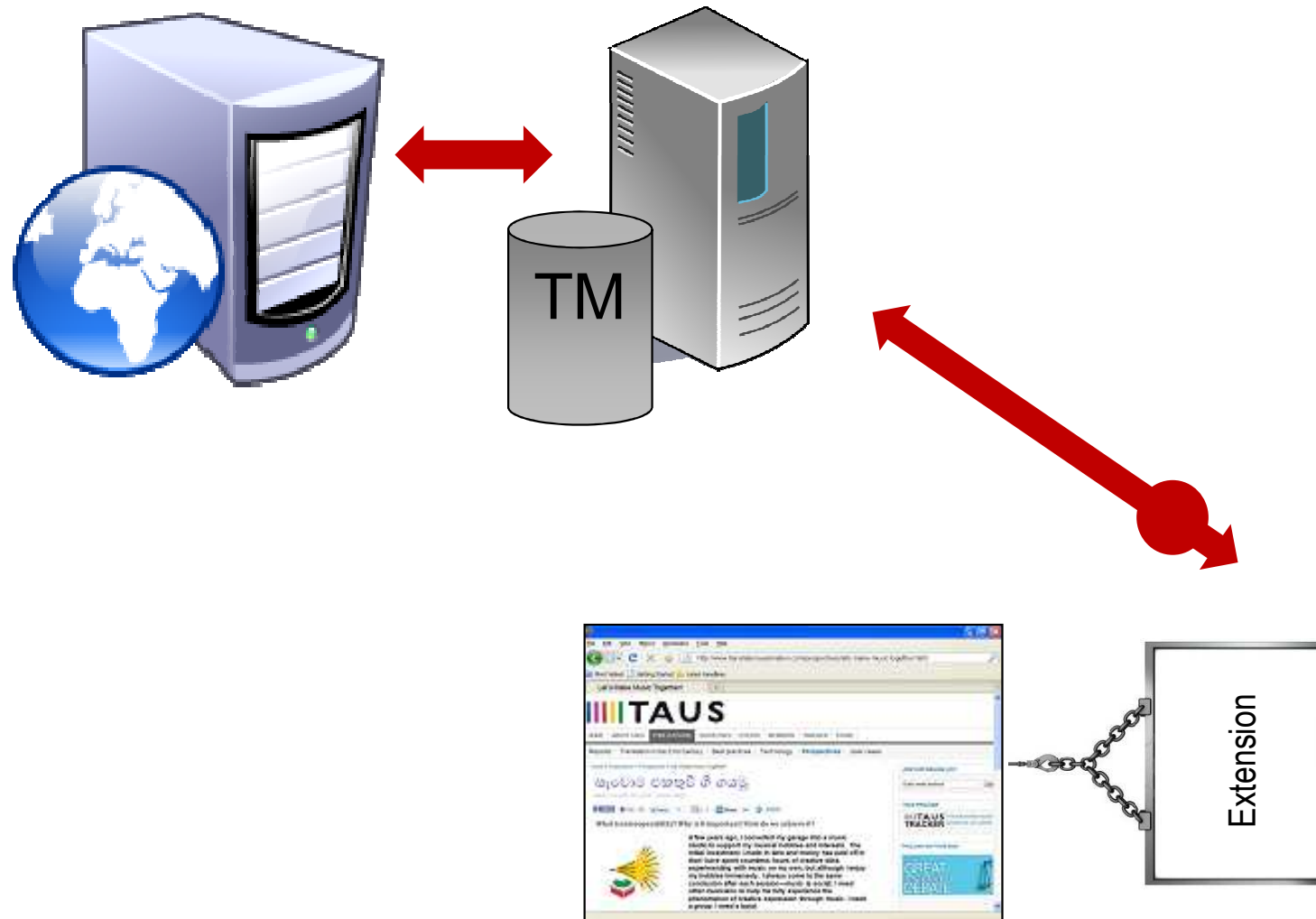
TAUS TRACKER

Find and shortlist machine solutions for your specific

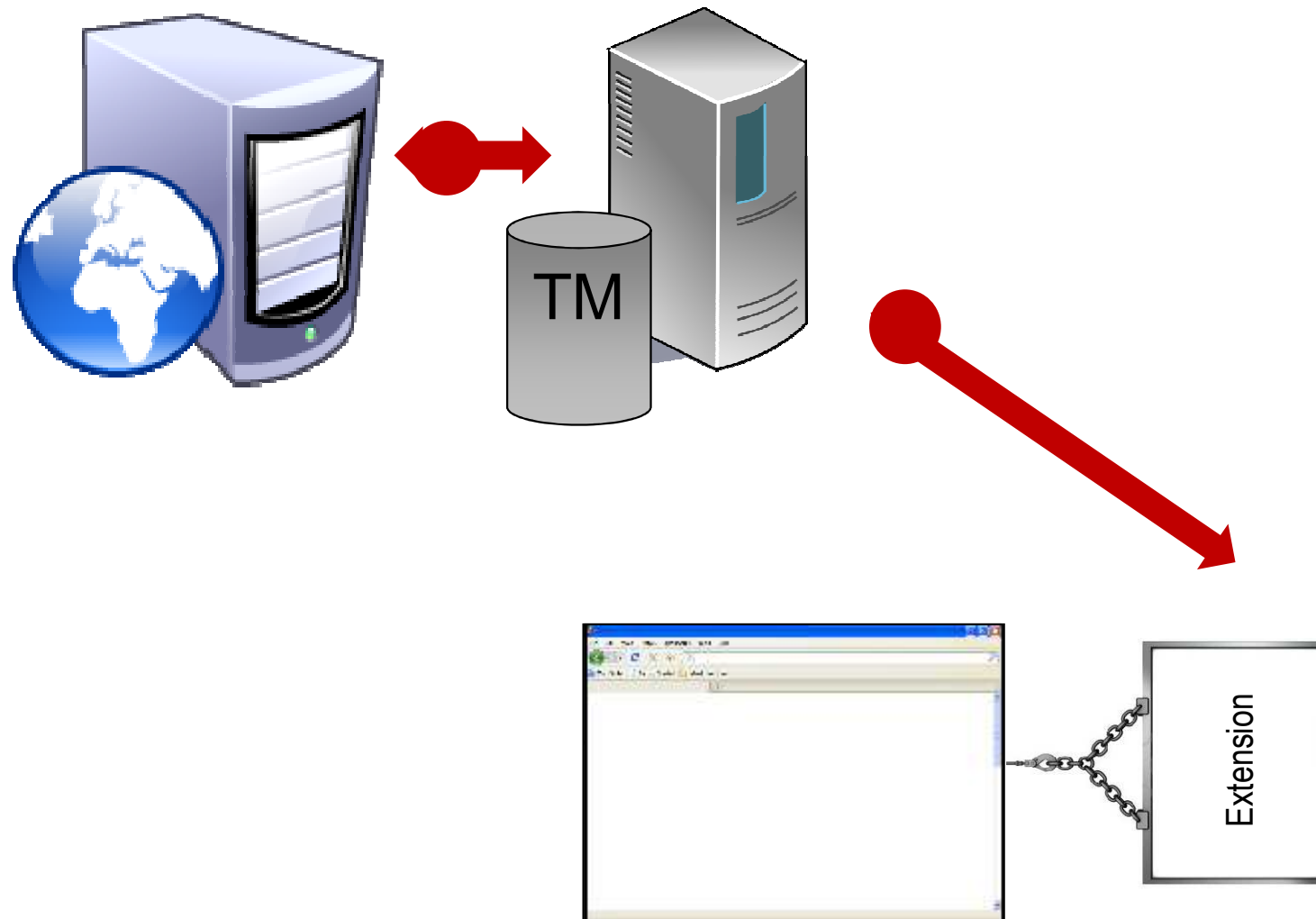
ProZ.com and TAUS host

THE GREAT TRANSLATION DEBATE






```
<xliff version='1.2'
  xmlns='urn:oasis:names:tc:xliff:document:1.2'>
<file original='site/e1/1223' source-language='en' target-language='si'
  datatype='plaintext'>
<body>
<trans-unit id='124i8j'>
  <source>Let's Make Music Together!</source>
  <target>සැවොම එකතු වී ගයමු </target>
  <context-group>
    <context context-type="element">
      xpath:html/body/div#ja-a-content/div#ja-clearfix/a
    </context>
  </context-group>
</trans-unit>
</body>
</file>
</xliff>
```





සැබවම එකතුවී ගියෙමු

FRIDAY, 08 APRIL 2011 09:52 | RUSTIN GIBBS

Share Like 14 Tweet 9 +1 0 Share 24

අන්තර් ප්‍රතිග්‍රහනය යනු කුමක් ද? එය වැදගත් වන්නේ කෙසේද?

A few years ago, I co-founded a music studio to support my initial investment in music. I realized that I have spent countless hours experimenting with my hobbies immensely. I reached a conclusion after each session that other musicians to help me understand the phenomenon of creative expression through music. I need a group. I need a band.



Open Link in New Tab
Open Link in New Window
Bookmark This Link
Save Link As...
Send Link...
Copy Link Location
Copy

edit: සැබවම එකතුවී ගියෙමු!

vote: සියලුම එකතුවී ගියෙමු! 46
අපි සැබවම තනු සෙවමු!
සැබවම එක්වී ගියෙමු! 472

JOIN OUR MAILING LIST

Enter email address

TAUS TRACKER

Find and shortlist machine solutions for your specific

ProZ.com and TAUS host

THE GREAT TRANSLATION DEBATE

Issues

- Legal issues
- L10n of non-textual content
- Constant updates
- Style/Formatting issues
- Translation voting issues
- Deployment issues

Benefits

- A practical web translation method for less-resourced languages
- Web site independent framework
- Non-intrusive mechanism for local language users to benefit from the vast knowledge embodied in the Internet
- A method to facilitate l10n of web pages (in early stage of l10n cycle) e.g. in context l10n./;
- TM based – more accurate translations
- Helps building of resources such as parallel corpora

Standardization considerations

- Different extension mechanisms



Standardization considerations

- Different rendering mechanisms/engines (esp. mobile)



Standardization Considerations

- Browser Extensions
- Rendering Engines and Fonts

Summery

- Multilingual Web is not just the top10 languages
- Browser extensions
 - Enable a majority to obtain the most out of the vast body of knowledge encompassed in the largely English
- Novel TM based web site localisation methodology
- Standardization considerations
 - Browser Extensions
 - Rendering Engines

Thank you