

Faculty Summit 2012

Riviera Maya, Mexico | May 23-25 | In partnership with CONACYT



Microsoft Translator Hub & the Mobile Travel Companion

Kristin M. Tolle, Ph.D. Director, Microsoft Research Connections

May 24th,2012

Statistical MT = Machine Learning Algorithms + Data

Algorithms

- A Decoder (the engine core)
- Multiple Translation Tables
- Multiple Target Language Models
- Efficient Training System
- Scalable, Reliable and Fast Runtime

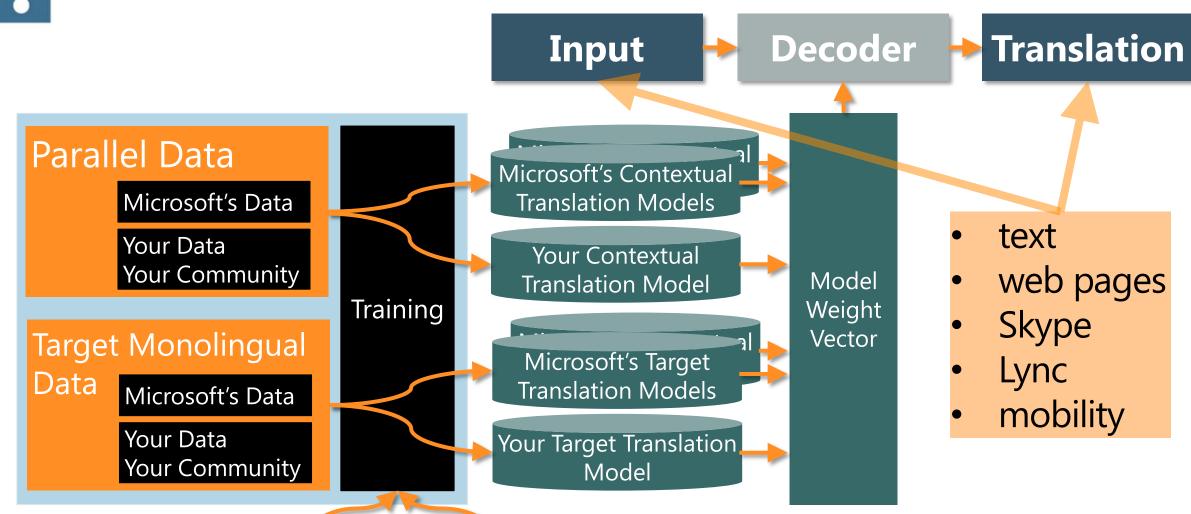
Data

- Parallel Documents
- Target Language Documents
- Domain-Specific Parallel Documents
- Domain-Specific Target Language
 Documents
- User edits, votes, ranking
- Tuning Set
- Test Set

Interesting New Developments



From Training to Runtime



Test Set



Microsoft^{*}

- - X

Search this site.

0

About Us | Schools | News | Employment

FOR PARENTS -

FOR STUDENTS -

FOR THE COMMUNITY

High-performing public schools

PARENTS

Supporting your student

STUDENTS

For success now and in the future

COMMUNITY

Strong community, strong schools



April Staff Star: Beth Weaver, teacher, McAuliffe Elementary

Whether she is encouraging students to be involved in their school ASB or serving as a faculty representative on the PTSA board, Beth Weaver is helping Christa McAuliffe Elementary students become "Future Ready."

Go to Staff Stars



Redmond High student receives National Merit Corporate-Sponsored Scholarship

Student among 1,000 distinguished students recognized nationwide

Lake Washington School District 04.18.2012 Teacher's Contract Extended, Calendar Set

School year ends one day later; secondary students have "transition"

Two LWSD Students Named National 04.04.2012 Achievement Scholarship Program Winners

Nine students from Washington state recognized

Upcoming Events

School Board Work Session: 05.07.2012 5 p.m. at the Resource Center

05.07.2012 School Board Meeting: 7 p.m. at the Resource Center

05.21.2012 School Board Meeting: 5:30 p.m. at the Resource

Center

05.21.2012 School Board Work Session: 7 p.m. at the Resource Center

05.25.2012 No School - LEAP Day

05.28.2012 No School - Memorial Day

06.04.2012 School Board Work Session: 5 p.m. at the Resource Center

What's New on this Site...

Secondary and elementary summer school catalogs now available

Mail-in registration begins May 14

Make online payments for fees

Convenience for junior high, high school parents

Need Assistance?

The Community Resources page has a list of community resources to support students, parents and families in our community.

How Do I...

Lagra mara hara

Register a student for school? Add money to a lunch account? Find which school a student will attend

Microsoft Research Faculty Riviera Maya, Mexico

Microsoft

04.19.2012



Latest News

04.19.2012 Redmond High student receives National Merit Corporate-Sponsored Scholarship

Student among 1,000 distinguished students recognized nationwide

04.18.2012 Lake Washington School District Teacher's Contract Extended, Calendar Set

School year ends one day later; secondary students have "transition"

Two LWSD Students Named National 04.04.2012 Achievement Scholarship Program Winners

Nine students from Washington state recognized

New STEM School to be housed 03.29.2012 temporarily at Eastlake High

02.28.2012

02.03.2012

Permitting delays push building opening back

Lake Washington School Board selects Dr. Traci Pierce as new superintendent

Experienced educational leader to continue current initiatives

Calendar changes due to snow days 02.23.2012

School year extended one day; students to attend May 4

Fourteen schools recognized with Washington Achievement Awards

Lake Washington High wins in three categories

PLUG LOAD

Upcoming Events

05.07.2012 School Board Work Session: 5 p.m. at the Resource Center

05.07.2012 School Board Meeting:

7 p.m. at the Resource Center

05.21.2012 School Board Meeting:

5:30 p.m. at the Resource

Center

05.21.2012 School Board Work Session: 7 p.m. at the Resource Center

05.25.2012 No School - LEAP Day

No School - Memorial Day 05.28.2012

06.04.2012 School Board Work Session:

5 p.m. at the Resource Center

To see all district events, see our online Tandem Calendar

What's New on this Site...

Secondary and elementary summer school catalogs now available Mail-in registration begins May 14

Make online payments for fees

Convenience for junior high, high school parents

Need Assistance?

The Community Resources page has a list of community resources to support students, parents and families in our community.

How Do I...

Register a student for school? Add money to a lunch account? Find which school a student will attend

Learn more here...

Choose task



Like us on Facebook Follow us on (witter

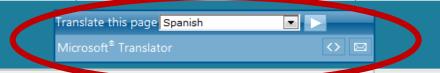
Sign up for LWSD Newsletters

Your email:

GO









Microsoft Research

Riviera Maya, Mexico | May



Latest News

Redmond High student receives 04.19.2012 National Merit Corporate-Sponsored Scholarship

Student among 1,000 distinguished students recognized nationwide

Lake Washington School District 04.18.2012 Teacher's Contract Extended, Calendar Set

School year ends one day later; secondary students have "transition"

Two LWSD Students Named National 04.04.2012 Achievement Scholarship Program Winners

Nine students from Washington state recognized

New STEM School to be housed 03.29.2012 temporarily at Eastlake High

Permitting delays push building opening back

Lake Washington School Board selects Dr. Traci Pierce as new

superintendent Experienced educational leader to continue current initiatives

Calendar changes due to snow days 02.23.2012

02.28.2012

02.03.2012

School year extended one day; students to attend May 4

Fourteen schools recognized with Washington Achievement Awards

Lake Washington High wins in three categories

Upcoming Events

05.07.2012 School Board Work Session:

5 p.m. at the Resource Center

05.07.2012 School Board Meeting:

7 p.m. at the Resource Center

School Board Meeting: 05.21.2012 5:30 p.m. at the Resource

Center

05.21.2012 School Board Work Session:

7 p.m. at the Resource Center

lline

05.25.2012 No School - LEAP Day

05.28.2012 No School - Memorial Day

06.04.2012 School Board Work Session:

Dutch Estonian Finnish French To see al German

Greek Haitian Creole

Hebrew Hindi

Hungarian Indonesian

Italian Japanese Korean Latvian Lithuanian Norwegian Polish

Portuguese Romanian Russian Slovak

Slovenian Spanish Swedish Thai

Ukrainian Vietnamese Translate this page Spanish

Turkish



Secondary and elementary summer school catalogs now available

Mail-in registration begins May 14

Make online payments for fees

Convenience for junior high, high school parents

Need Assistance?

The Community Resources page has a list of community resources to support students. parents and families in our community.

How Do I...

Register a student for school? Add money to a lunch account? Find which school a student will attend

Learn more here...

Choose task



Follow us on Cwilter

Sign up for LWSD Newsletters

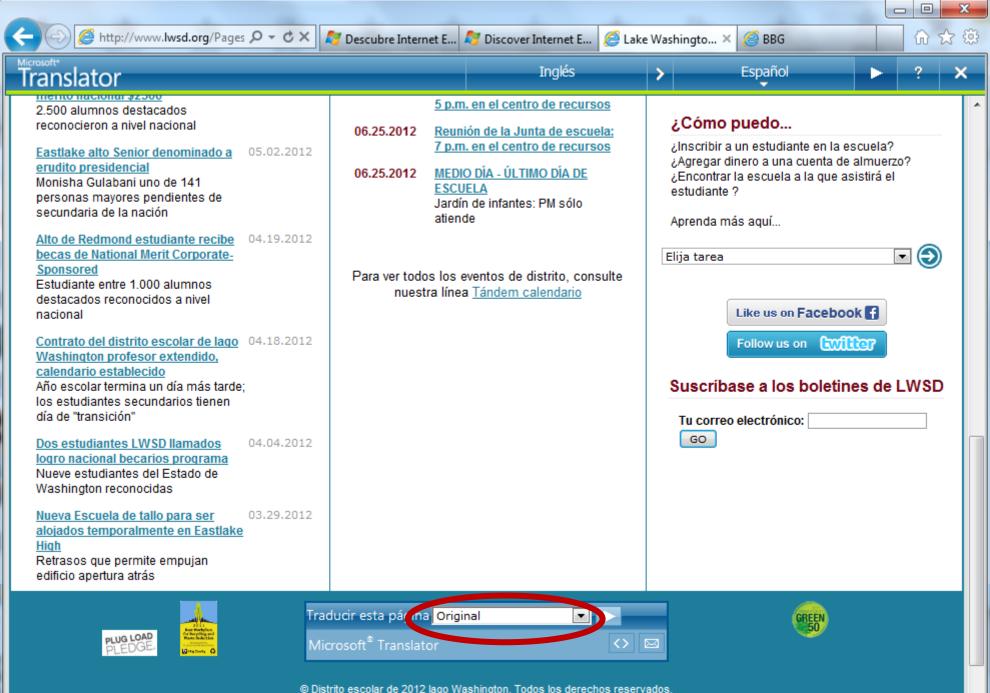
Your email: GO



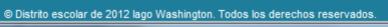






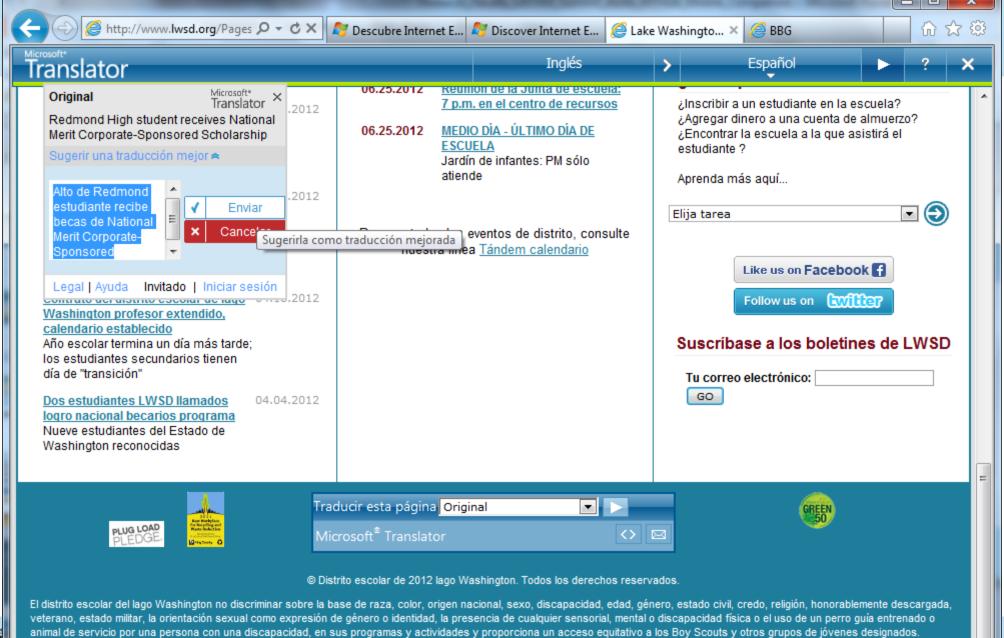


Microsoft Res Facu Riviera Maya,









Facul

Microsoft

Declaración de política completa



Community Results: LWSD

Lake Washington School District currently has about 30 volunteer parents.

Arabic (1)

Chinese – Simplified (4)

Danish (1)

French (2)

German (3)

Hindi (2)

Italian (1)

Japanese (1)

Polish (1)

Portuguese (1)

Russian (5)

Spanish (7)

Thai (1)

Turkish (1)

- Reward editors
 - API counts edits and maintains leaderboard
- No spamming has happened to date





Jueves 17 de mayo de 2012.





BUSCAR Escriba lo que busca... Q

01 800 822-2628 Lada sin Costo

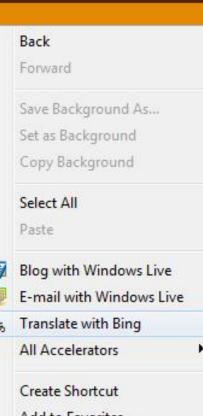
Contacto



INICIO

Nuestra Universidad





Add to Favorites... View Source Encoding Print... Print Preview... Refresh

Curso de Buceo

Participan estudiantes en Congreso de

Vida Universitaria

Auto-Detect

> Eventos Anáhuac

Noticias

Celebración por el Día del Maestro

Servicios

English

Translate this page

Powered by Microsoft® Translator

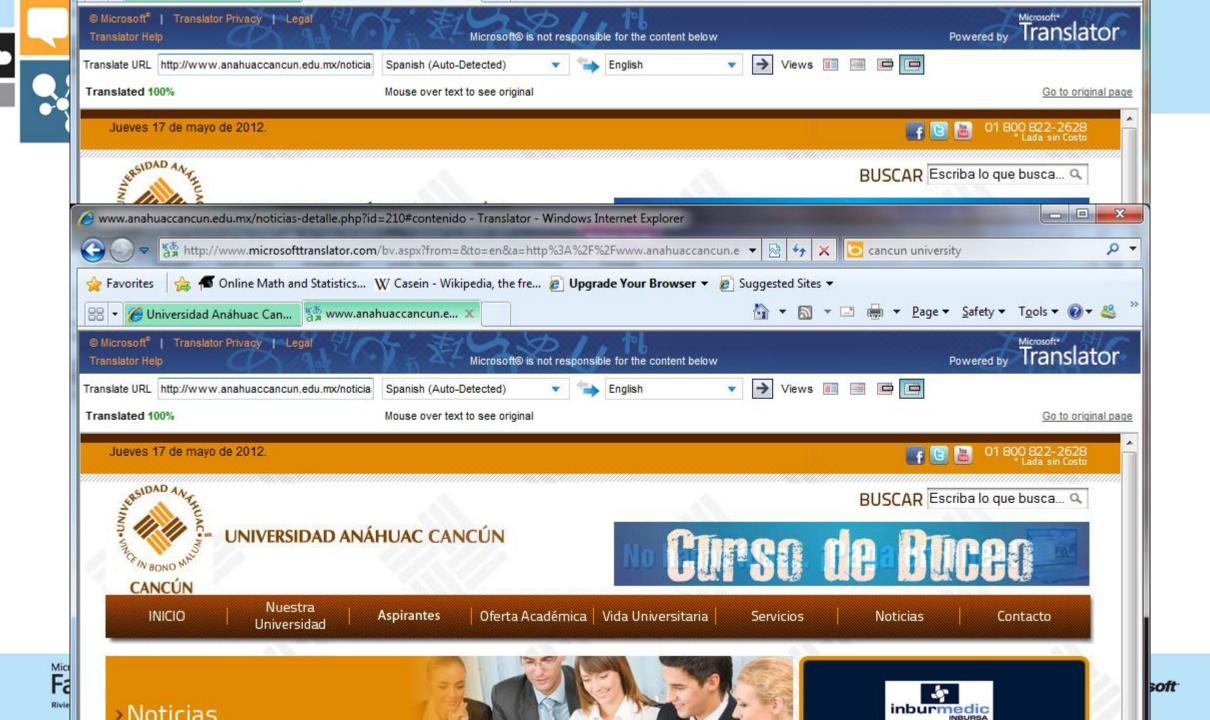
ión docente para May rsidad.

dencias y nuevos May

ción de la Escuela Jun de la Escuela de 14

Faculty Summit 2012 Riviera Maya, Mexico | May 23-25 | In partnership with CONACYT

Microsoft



What about languages that are not yet supported?

La eliminación definitiva de la libertad de expresión fue la pérdida de una lengua - John Ralston Saul

Translator Hub

http://hub.microsofttranslator.com

- Enables <u>companies</u>, <u>language communities</u> and <u>educators</u> to build <u>custom</u>, <u>private</u>, <u>domain specific</u>, <u>web accessible</u> automatic translation models.
- Empowers users to build translation systems independently.
- Free web service application that leverages Microsoft Translator's Machine Learning algorithms and massive language data sources.
- APIs and guides for developers SOAP, AJAX, HTTP, and OData





Working with Communities to Build Models

come.

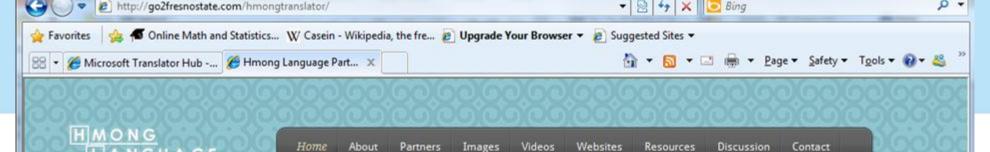


Dr. Phong Yang CSU Fresno, Linguistics Dept. Hmong Community Leader

Hmong Daw:
Released on <u>UNESCO's Intl.</u>
Mother Languages Day
Working to preserve their
language for generations to







LANGUAGE PARTNERS

Discussion



ABOUT

The Hmong Translator Project started in November 2011 and is a collaborative effort between the Microsoft Research team and the Hmong community.



PARTNERS

The Hmong Language Partners is a community of individuals, organizations, and entities who want to preserve the Hmong language.



NEWS

Find out about the latest development on this project. Give us your feedback, comments and suggestions to help improve this tool.



BECOME A PARTNER

We are looking to expand our list of partners and participants. If your organization would like to take part in the development of this Hmong translator,

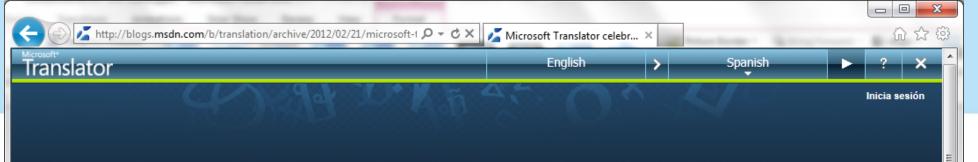
"Of the approximately 7,000 languages in the world, less than 100 of them have automated instant translation capabilities."

TRANSLATE









Blogs MSDN > Blog del equipo oficial de traductor de Microsoft (y Bing Translator) > Traductor de Microsoft celebra día internacional de la madre de la lengua con el lanzamiento de Hmong

Traductor de Microsoft celebra día internacional de la madre de la lengua con el lanzamiento de Hmong



Microsoft Translator 21 De febrero de 2012 7:46 AM



En la celebración del Día internacional de la madre de la lengua, nos complace anunciar la incorporación de la lengua Hmong a nuestra lista de los idiomas admitidos, hecha posible por una estrecha colaboración con la comunidad Hmong. Cualquiera puede probar el nuevo idioma en el sitio de Bing Translator o llamar a través del servicio de web de Microsoft Translator (Hmong Daw, mww de código de idioma).

Hmong Daw es el dialecto de la etnia Hmong el sistema admite, también conocido como blanco Hmong.

Instrumentales para este esfuerzo fueron los miembros de la Comunidad Hmong, fueron capaces de aprovechar las nuevas herramientas de traductor de Microsoft para ayudar a preservar y revitalizar su idioma en línea. Estas nuevas herramientas. actualmente en fase beta, habilitar soporte de traducción automática



de idiomas adicionales, o sistemas de calidad superiores de edificio de terminología específica y estilo en los idiomas establecidos.

La adición de la lengua de la etnia Hmong es un ejemplo de la primera hipótesis: los miembros de la comunidad utilizados existentes traducción materiales y nuevas características de Microsoft Translator para formar un nuevo motor de traducción. Esto aprovecha capacidades de aprendizaje de Microsoft Translator, que pueden aprender cómo traducir de un conjunto de documentos paralelos (mismo documento en dos idiomas), diccionarios y textos en el idioma a traducir a (Hmong en este caso). Además de enseñar el motor de un nuevo idioma, ellos también participan a miembros de la Comunidad, socios y colaboradores para crear y revisar versiones mejoradas del sistema de traducción automática y recopilar comentarios cualitativos acerca de cada sistema "capacitado". Implementación de un sistema que alcanza un cierto nivel de calidad permite su uso transparente con las API estándar de traductor de Microsoft y muchos escenarios impulsados por la API, como el widget de traducción de web. Comentarios que se generaron a través de estos escenarios pueden utilizarse de nuevo en el proceso de formación, creando un bucle virtuoso para mejorar la calidad de la traducción. Permaneced atentos para más



Tareas comunes

- Inicio de blog
- RSS para comentarios
- RSS para puestos

Búsqueda



Mensajes más recientes

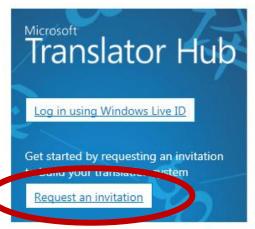
A audazmente recuperaremos Going places con la app de traductor más innovadora para Windows Phone

Traductor de Microsoft celebra día internacional de la madre de la lengua con el lanzamiento de Hmong

Principales sitios va Global con traductor API

API de traductor de Microsoft ahora soporta OData, detalles de mayores volúmenes anunciados los precios





Where language meets the world

Bridging languages, cultures and technology



Microsoft Translator Hub is helping smaller languages thrive by putting the power to build machine translation systems in the hands of local communities.

- K. David Harrison, National Geographic Fellow, author and linguist

MORE INFORMATION

Microsoft Translator Hub Overview

Microsoft Translator Hub Forum

ABOUT MICROSOFT TRANSLATOR HUB

Microsoft Translator Hub empowers businesses and communities to build, improve, and deploy customized automatic language translation systems—bringing better and specialized translation quality to established languages, as well as the many native languages of the world that are not yet supported by major translation providers.

Powered by Windows Azure, Microsoft Translator Hub is an extension of the Microsoft Translator platform and service. You can build a superior translation system easily, within a private website, by combining your translated documents with the power of Microsoft Translator's big data back end. Once you are satisfied with your translation, you may share it publically on the web.









VIDEOS







Projects Community Help

Microsoft^{*} Translator Hub

Create Project

A project allows you to build a new translation system between two languages using a terminology & style of your choice.

View Help

Project Name:*

Description:

Mayan Preservation Project

Local Yucatan Mayan Langauge Project

Source Language:*

Target Language:*

Category Descriptor:

Category:

Maya, Mopán

*My source language is not listed. Request a language to be added >

*My target language is not listed. Request a language to be added >

▼ 0

General

English

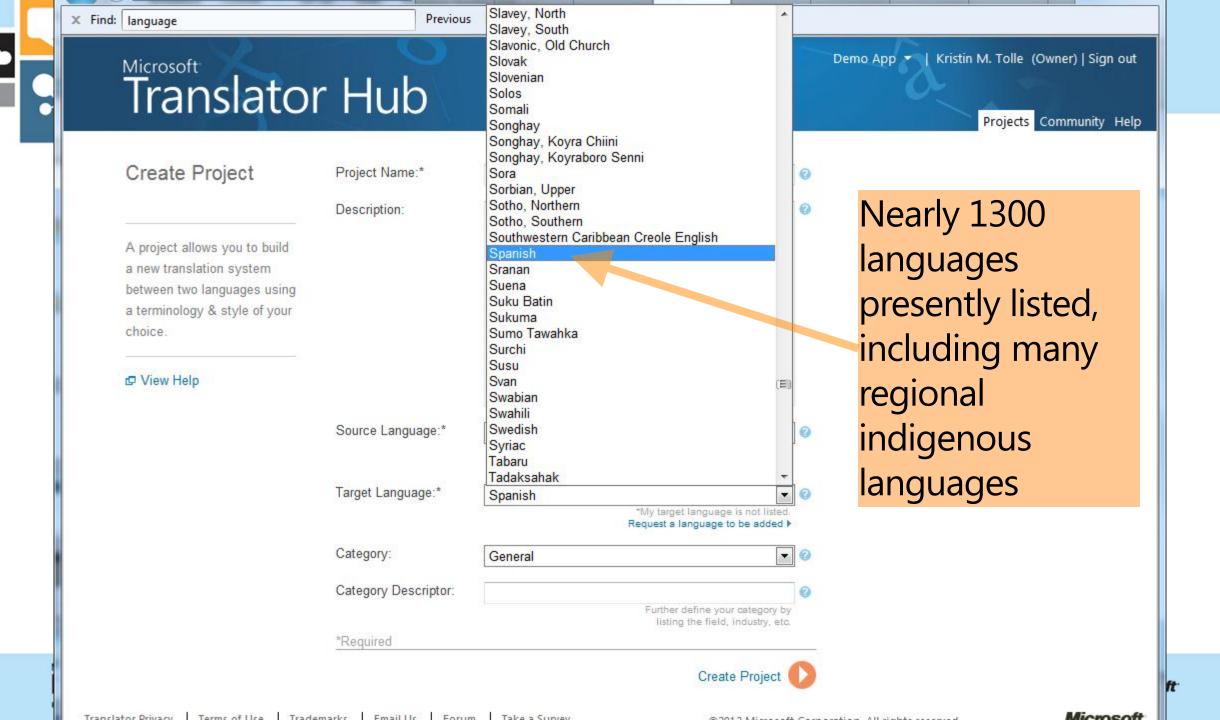
Further define your category by listing the field, industry, etc.

*Required

Create Project

Manage Projects Manage Documents Create Project

Request Language Reviewer Dashboard



Microsoft¹ Translator Hub

Create Project

A project allows you to build a new translation system between two languages using a terminology & style of your choice.

View Help

Project Name:*

Description:

Fine Arts Food Geography, Anthropology Government

Energy, Water and Utilities

Healthcare History

Home & Garden

Internet Language

General

Agriculture Animals

Automotive Beauty

Business Chemicals Clothing

Custom Education |

Electronics

Financials

Arts & Entertainment

Law

Literature Medicine

Military Science

Music Philosophy Political Science

General

Category Descriptor:

Source Language:*

Target Language:*

*Required

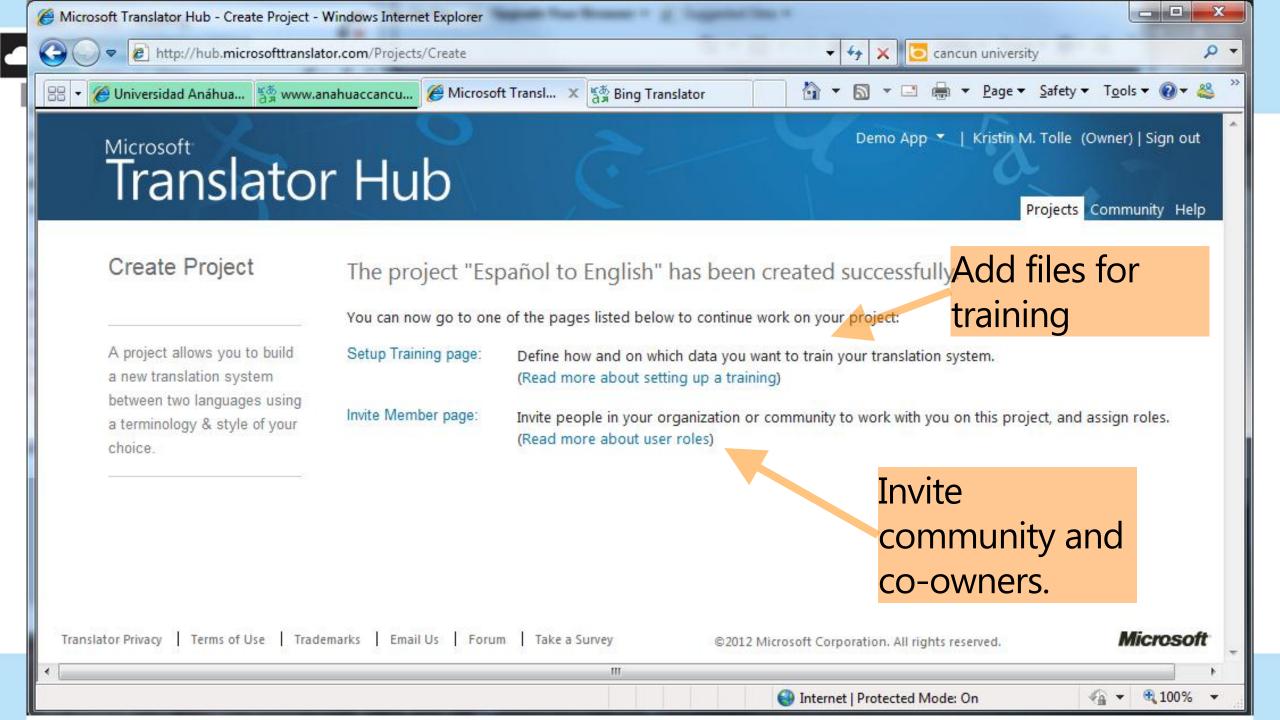
Category:

For developing custom, domain specific models.

Great for education and socially relevant applications.

Further define your category by listing the field, industry, etc.

Microsoft^{*}



Projects Community Help

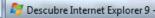
Setup Training Español to English Language Pair: Spanish to English | Category: General Related Tasks Training Name:* Training #1 Define how your custom Save this training V 0 Use Microsoft models: translation system should be trained & tested. Search training documents.. Training Tuning (Auto) Testing (Auto) Show All Show Selected [add documents **Setup Training: Upload Documents** X nce Count 🕝 Note: 10 Select a document or an archive file (ZIP, GZ, TGZ) containing documents that you want to upload. **Document Naming Style:** 20 <document name>.TMX OR 20 <document name> [esn | enu].[TXT | HTML | HTM | PDF | DOCX | ALIGN] "esn" indicates that the document contains "Spanish" sentences and 20 "enu" indicates that the document contains "English" sentences 20 e.g: Doc1_esn.txt, Doc1_enu.txt, Doc2_enu.pdf, Doc3.tmx, Doc4_enu.docx 20 Maximum Size: 50MB. \\er-data\mt\TestData\es-es to en-us\oov combined esn.zip Browse... 20 By uploading my documents, I confirm that the content I submit does not infringe the copyrights, publicity rights, privacy rights or other intellectual property rights of others. I further confirm that I have the sufficient rights in the 20 content to grant Microsoft the licenses provided in the Terms of Use. 20 Upload Cancel

If documents with the same name exist, allow overwrite.

MIND IN THE HIGH CHIEF AND A

20







Market Microsof... P - C X

Página principal de Internet Explorer Características

Explorar Windows

Descubre Internet Explorer 9



Comprar

Comparar exploradores

Buscar en este sitio web

Productos

Sugerencias

Descargas



Kristin Tolle, Cerrar sesión | México | Todos los sitios de Microsoft

Ayuda y procedimientos



Search this website

Explore Windows Products Downloads Shop Videos Internet Explorer home Features Compare browsers

Discover Internet Explorer 9

Internet Explorer 9 features

Fast

Performance improvements

Clean

New Tab page

Notification Bar

One Box

Pinned Sites

Tab enhancements

Web-focused browsing

Windows 7 integration

Trusted

ActiveX Filtering

Add-on Performance Advisor

Compatibility View

Cross-site scripting filter

Domain highlighting

Download Manager

InPrivate Browsing

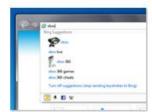
SmartScreen Filter

A more beautiful wel



All-around fast

Fast and responsive web experiences for Windows.



One Box

Go to a web address or start a search—all from the same place.

Características de Una nueva forma de disfrutar la Web

Rápido Mejoras en el rendimiento

Internet Explorer 9

Sencillo

Windows

Página principal

Página Nueva pestaña

Barra de notificación

One Box

Sitios anclados

Mejoras de pestañas

Navegación centrada en Internet

Integración con Windows 7

Seguridad

Filtrado ActiveX

Asesor de rendimiento de complementos

Vista de compatibilidad

Filtro de scripts de sitios

Resaltado de dominios



Rapidez en todo

Una experiencia web más rápida y con mejor respuesta para Windows.



Centrado en tus sitios web

El nuevo navegador hace posible que los sitios web desplieguen todos sus encantos.



Perfecta integración con Windows 7

Sitios anclados, Jump Lists y más: todo funciona conjuntamente.



One Box

Ve a una dirección web o inicia



Sitios anclados

Accede a tus sitios favoritos con



Filtro SmartScreen

Protección integrada para





Microsoft¹ Translator Hub

Demo App ▼ | Kristin Tolle (Owner) | Sign out

Projects Community Help

View Training Results

See the training results, review translations for accuracy, clone it to modify & resubmit or deploy your translation system.

View Help

Español to English

Language Pair: Spanish to English | Category: General



Related Tasks

- · Review translations
- · Clone this training
- · Test translations

Training (Auto)	Testing (Auto)
saming (mana)	January Comment

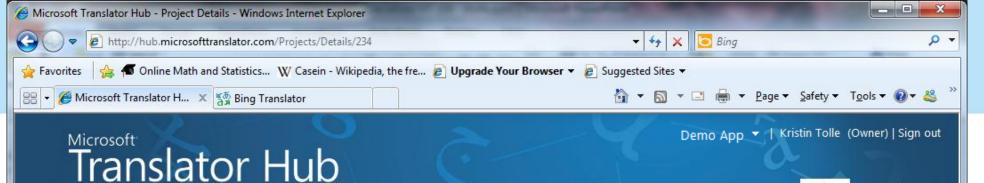
Search training documents...

Total aligned sentence count: 47,742 | Total extracted sentence count: 837,868

	Name	Type 🔺	Extracted Sentence Count ©	Aligned Sentence Count (2)	
006	1.pdf	Parallel	36 / 35	26	
03_	387.pdf	Parallel	68 / 66	62	
05_	bw_broch.pdf	Parallel	33 / 33	29	=
□ 05b	roch.pdf	Parallel	33 / 33	31	
07_	042b.pdf	Parallel	60 / 60	60	
09_	891a.pdf	Parallel	40 / 36	23	
090	4MN_HCBOR.pdf	Parallel	39 / 28	14	
11_	055.pdf	Parallel	19 / 18	14	
117	78.pdf	Parallel	70 / 71	66	
135	80-P_DWSD.pdf	Parallel	16 / 14	9	
1 4	001.pdf	Parallel	147 / 144	125	Ŧ







Project Details

See list of trainings, browse to see training details or add a training.

☑ View Help

Español to English

Language Pair: Spanish to English | Category: General

Training

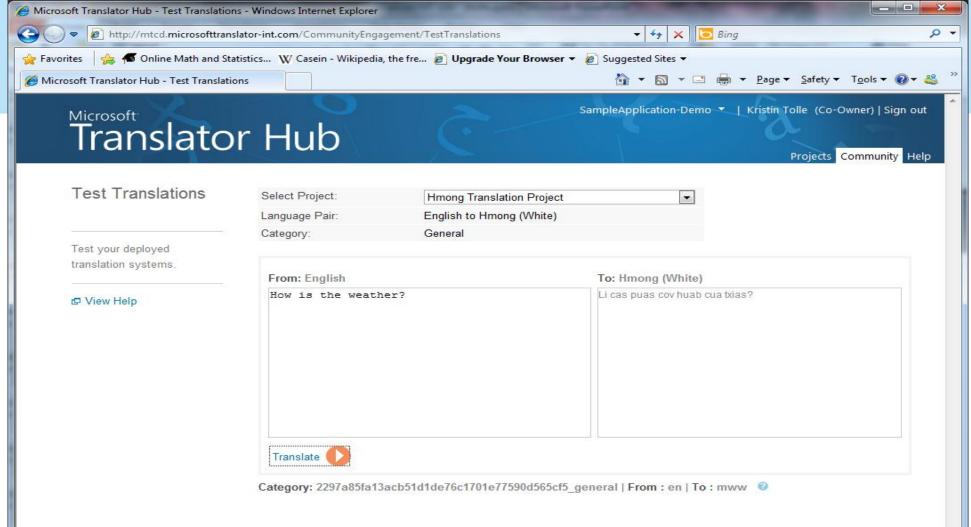
+ Add Training

Projects Community Help

Name	Status	BLEU	Aligned Sentence Count		Monolingual	Uses Microsoft	
			Training	Tuning	Testing	Sentence Count	Model
Traini	Succeeded	30.13	40,318	2,334	2,207	762,830	Yes
Traini	Deployed	30.97	47,742	2,334	2,207	757,026	Yes
Traini	Drafted	N/A	0	0	0	757,026	Yes
Traini	Drafted	N/A	0	0	0	193,777	Yes



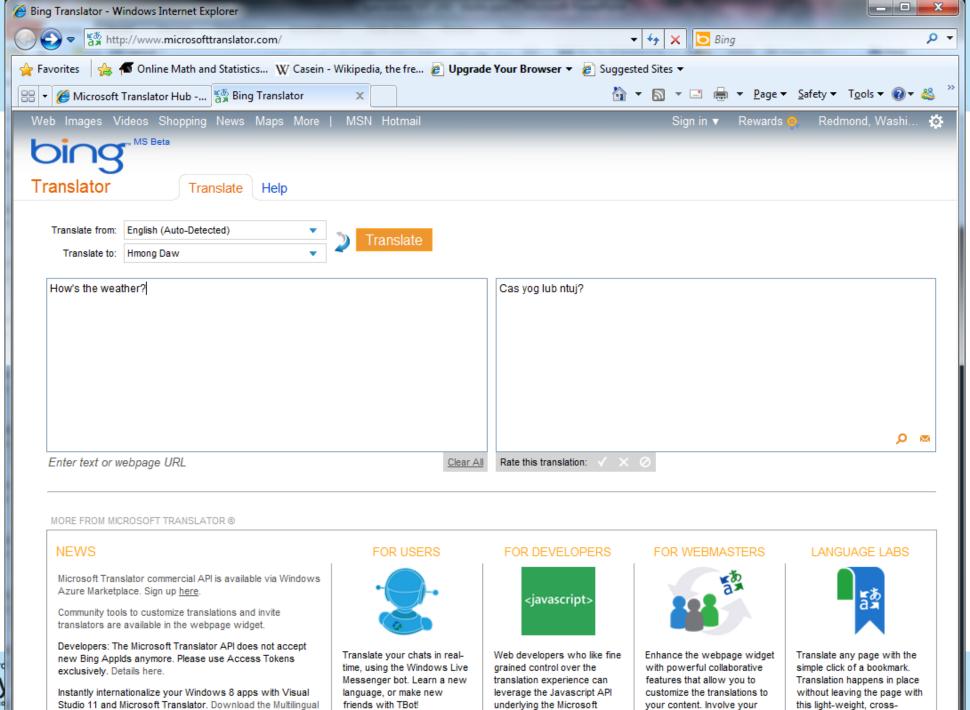












Translator widest

Microsoft

browner plugin



Translator Hub

Projects Community Help

Review Document

Review Hmong (White) translations of English sentences and submit alternate translations by hovering the mouse over the Hmong (White) sentence that can be improved.

View Help

File: 000_news1.txt

Language Pair: English to Hmong (White) | Category: General

Shared on Friday, March 30, 2012 1:24 AM by Rahul Sharma
Page

English Sentence	Hmong (White) Sentence
Ex-England keepers blast plans to move goalposts LONDON, Jan 3 (AFP) $$	Cov neeg khawscia Askiv-qub rhoob cov kev npajsiab los tshem cov ncejhomphiaj LONDON, Ibhlis 3 (AFP)
8 percent on the day to end at a record-high 3,715.	8 feem pua saum qhov hnub los xaus ntawm if qhov teevtsegsiab 3,715.
Gilts were off by an average of a quarter-point, again suffering from uncertainties about the government's future.	Gilts twb yog tawmnrauv losntawm ib qhov pesnrab ntawm ib Original Microsoft* Translator × bntawm cov kev tsis ib. Ganzuri will be 63 on January 12.
"President Mubarak has approved the formation of the new government," said Ganzuri.	Provide a better translation a hov txoosseem ri.
Ganzuri will be 63 on January 12. The Togalese National Assembly rejected Wednesday the budget presented by the government of Edem Kodjo.	Ganzuri yuav muaj 63 xyoos lub ib hlis tim 12.
Justice: Faruk Seif al-Nasr	
Housing: Mohamed Ibrahim Suleiman "Major markets hit their highs last year on the strength of Wall Street, but Tokyo had been lagging behind."	Ganzuri yuav muaj (muaj[have]) 63 xyoos lub ib hlis tim 12. ev siab xyoo kawg tabs s Tokyo twb
The Suara Karya daily reported that over 10,000 structures in East and North Aceh had been damaged or destroyed.	Ganzuri yuav yog 63 saum ibhlis 2. https://doi.org/10.000/ceh tau yog
	Dashboard Kristin Sign out





Data is the Key to Machine Translation Quality

Collaborative Translation Store

Microsoft Translator API Collaborative functions

Web site widget with Collaboration

Community inputs from the Microsoft Translator Hub

Collaborative Translation Memory Runtime

100% match at runtime

Trained modelsParallel training data
Your target language data

- Immediate reward
- Immediate credit in leaderboard
- Immediate validation and ranking





Collaboration is the Key to Language Vitality

- User's parallel data combined with community feedback are excellent sources of in-domain training data
- Achieve good quality by combining it with Microsoft's big data.
- Microsoft Translator Widget collects data from your community
- · Microsoft Translator Hub collects data from you and your community to train the <u>best system possible</u>.



Your Mobile Travel Companion

Translation of...

- Speech
- Live Video
- Text
- Images

using...

- Automatic Translation
- Human Aided Translation
 - via your social network

with smarts...

location and context

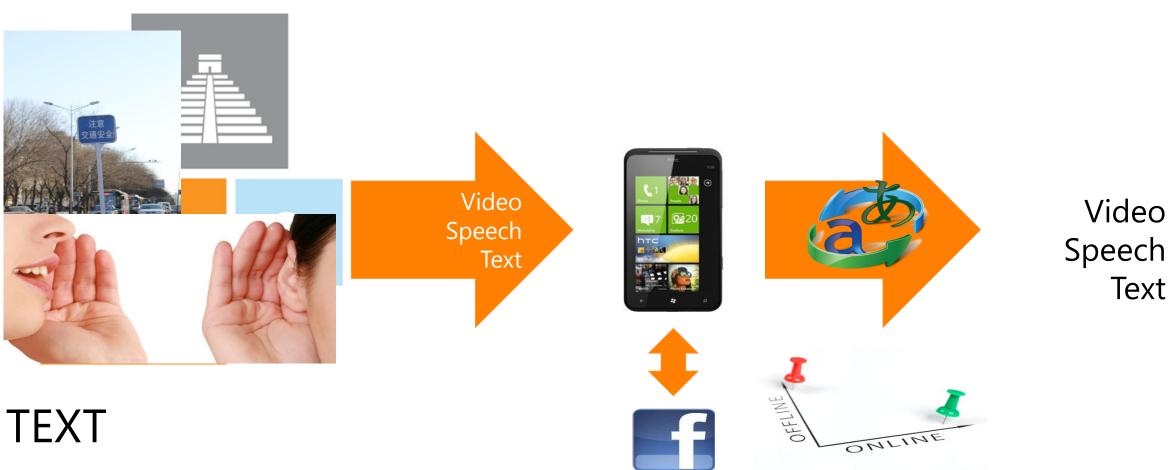
implemented as a...

Windows Phone App



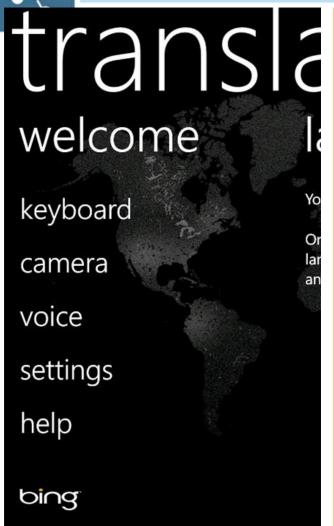
Mobile Travel Companion

Your Language Navigator in Foreign Lands



Video

Translator v2 for Windows Phone



- Human-quality (offline) phrasebooks
- Trusted translations via user's social graph
- Smart features using location & context
- Supports speech and camera input/output



KEYBOARD

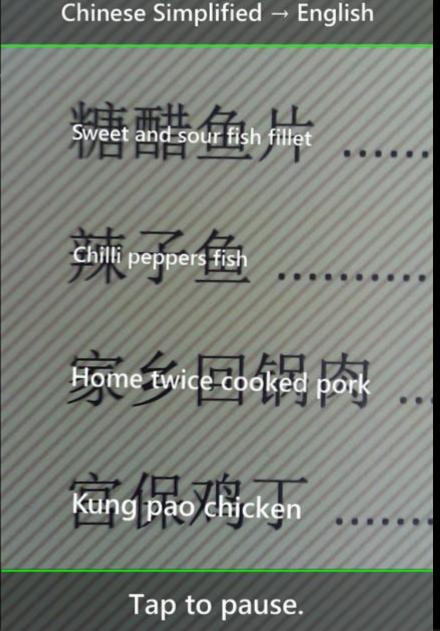
type alternativ

How are you?

English → Spanish



¿Cómo estás? How are you?



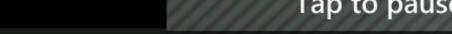
VOICE

speak alternativ

English (US) → Spanish



listen



• • •





...





anslato

TRANSLATOR

what's new all l

LANGUAGES

details



Translate text and images between French and English when your phone is offline.

FRENCH LANGUAGE OFFLINE

French

English

This download may take a few minutes. To minimize data usage, we recommend you use a WiFi connection.

Size: 35682 KB

Downloading...

languages



French Language Offline

French = English



Spanish Language Offline Spanish

English



French Language Offline French = English



Italian Language Offline Italian

English



German Language Offline German

English



Chinese Language Offline



view all

_ 0 X



search apps and games

Discover Buy Marketplace How-to My Phone

Sign in →

United States

Apps Games

travel + navigation / language

Translator



★★★★ Ratings: 1053







Use your camera to translate printed text, voice recognition to translate spoken text, or the keyboard to enter text that you want to translate.

Use your phone's camera to make translating signs, menus, newspapers, or any printed text a snap. Just aim the camera at the text and see the translation in secds.

Download languages and phrase books so you can translate text and images when your phone isn't connected to a network.

Speak a word or phrase that you want translated and hear the translations spoken in a native speaker's accent. Voice translations require a network connection.

Improve your vocabulary with the Word of the Day on the Start screen.

When you're traveling, need to overcome a language barrier, or trying to learn a foreign language, Translator is your go-to app.









anslato







Descubrir Comprar Marketplace Instrucciones Mi teléfono

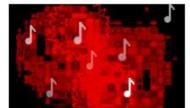
Kristin Tolle *

D

Uso compartido más fácil, aplicaciones más inteligentes y una mejor experiencia web













Questions??

Microsoft®