



Perspectives from the world of technology

British-Irish Council IML Languages Conference

9 November 2015

Laura Toby,
Localisation Vendor Manager, Google

Agenda

The Google philosophy

Languages at Google

Minority language initiatives

Language communities & the role of technology

GOOGLE

IMPACT



A group of five students are standing in front of a modern glass building with the Google logo. The students are dressed in casual attire, including a plaid shirt, a yellow top, a grey t-shirt with the Google logo, a white shirt, and a grey hoodie. They are engaged in conversation. A yellow and red bicycle with a basket is parked in the foreground. The scene is set outdoors with green trees and a clear sky.

Google

CULTURE



SCALE

The Google philosophy....

Fast is better than slow

Follow the user and all else will follow

Don't be evil

The need for information crosses all borders

Languages at Google

Google localisation: global & local

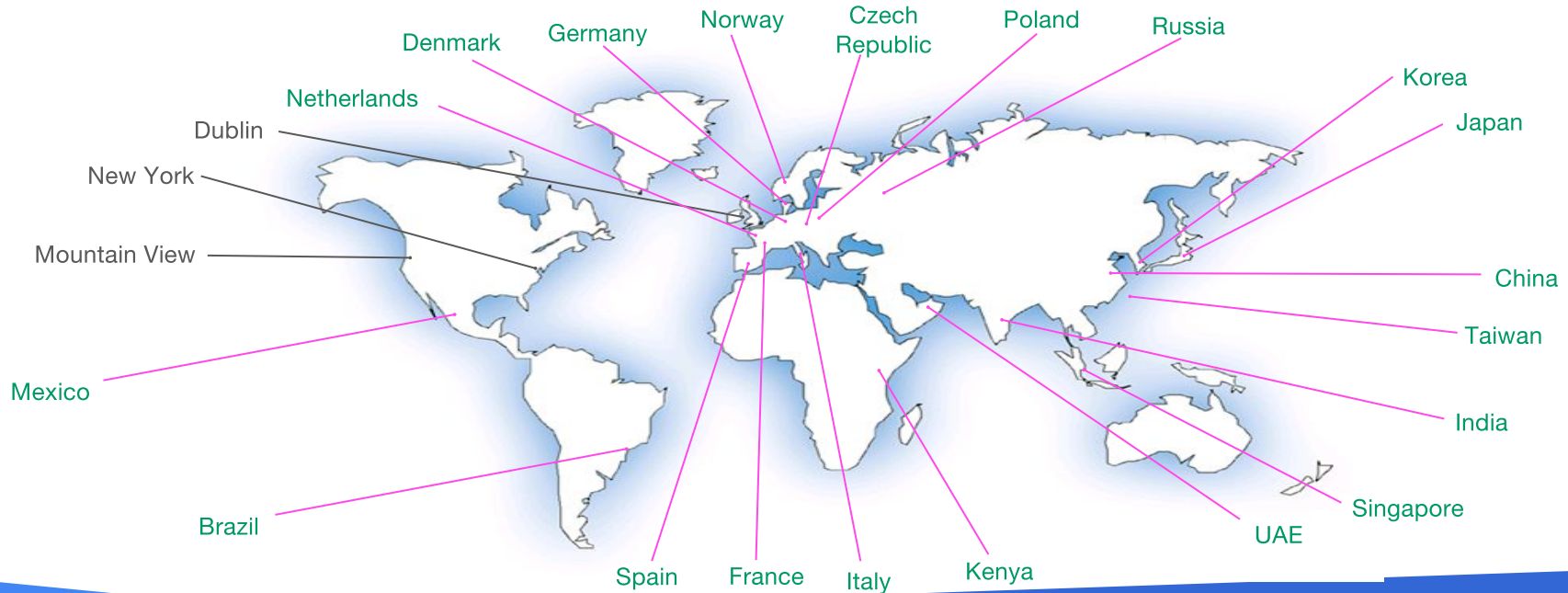
Proprietary + Confidential

In-country Language Specialists for top languages

Localization PMs in Mountain View, New York, Dublin

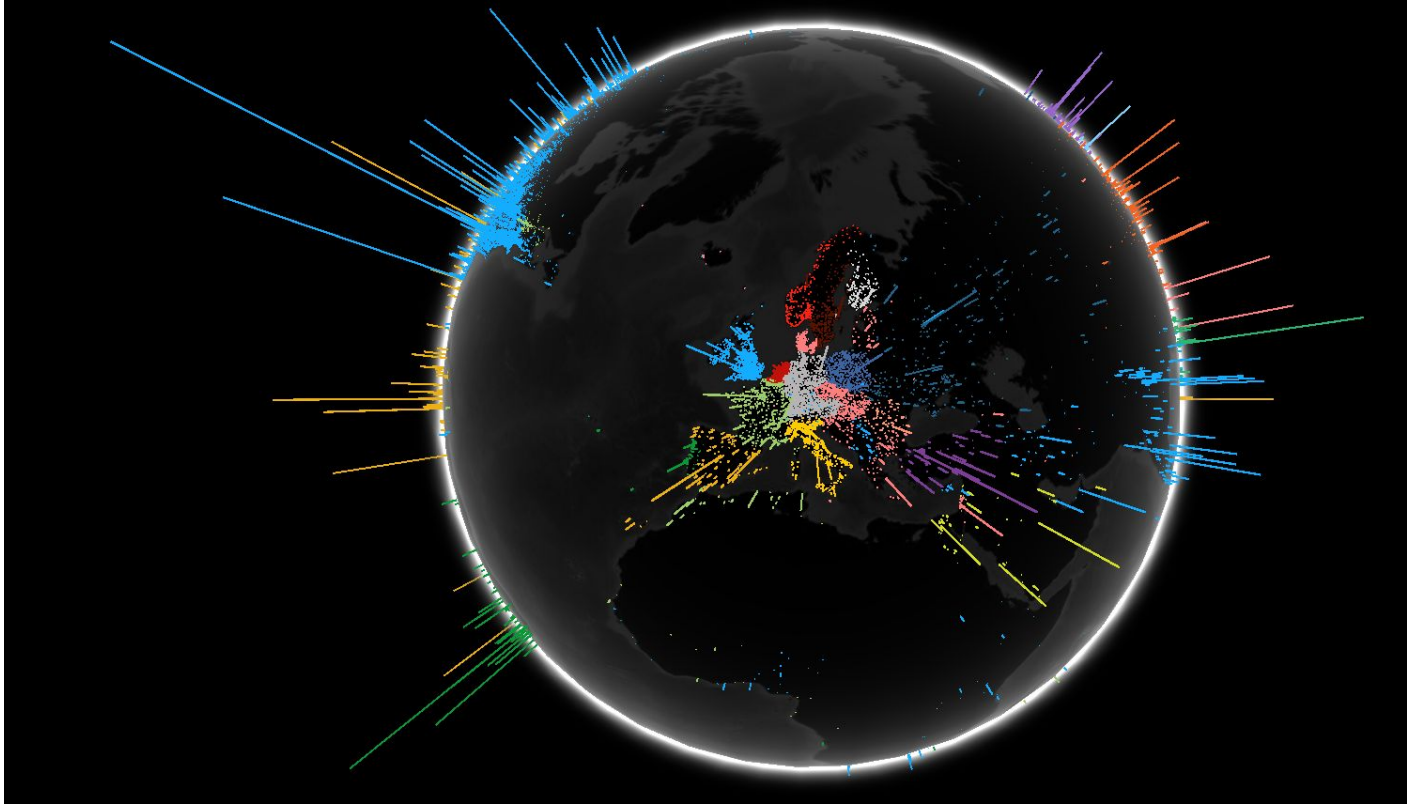
Localization Operations in Mountain View and Dublin

~60 people



... into 79+ languages...

Proprietary + Confidential



Localisation @Scale

Proprietary + Confidential

Over **7000** languages globally: any company can't support all

Goal: enable people worldwide to access info via one of the languages available

- Google participates in initiatives like **Unicode**: ensure anyone can write & communicate and develop content in their language (text support)
- Focus on **support for languages**:
 - Keyboards via Google [input tools](#) so anyone can create new content, chat, email, etc. in any language
 - Publicly available [fonts](#) for all characters in all written languages
 - [Google Translator Toolkit](#): free translation tool available to everyone

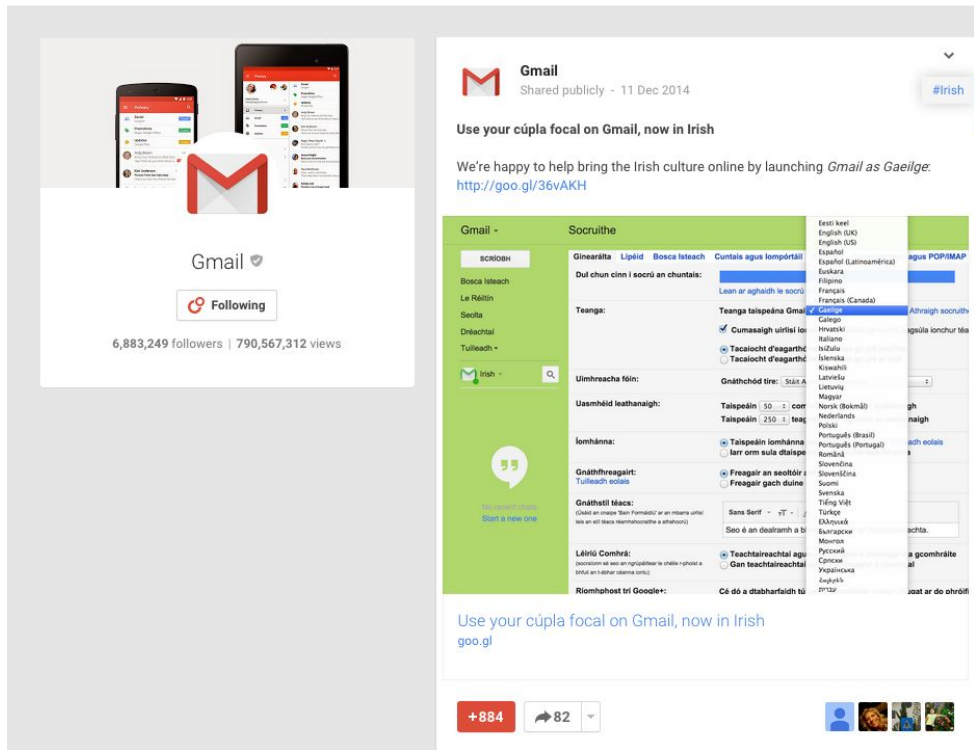
Minority Language Initiatives

Gmail: volunteer translations for Welsh, Cherokee & Irish

20% projects

Googlers in partnership with:

- Cherokee Nation
- Welsh Language Commission
- Prof Kevin Scannell



Google Translate

Proprietary + Confidential

90+ languages

- Corporates from language bodies & institutions help provide this service
- Individuals & groups can contribute

Most recent: Frisian

Google Translate

Your world.
Now with 90 languages.

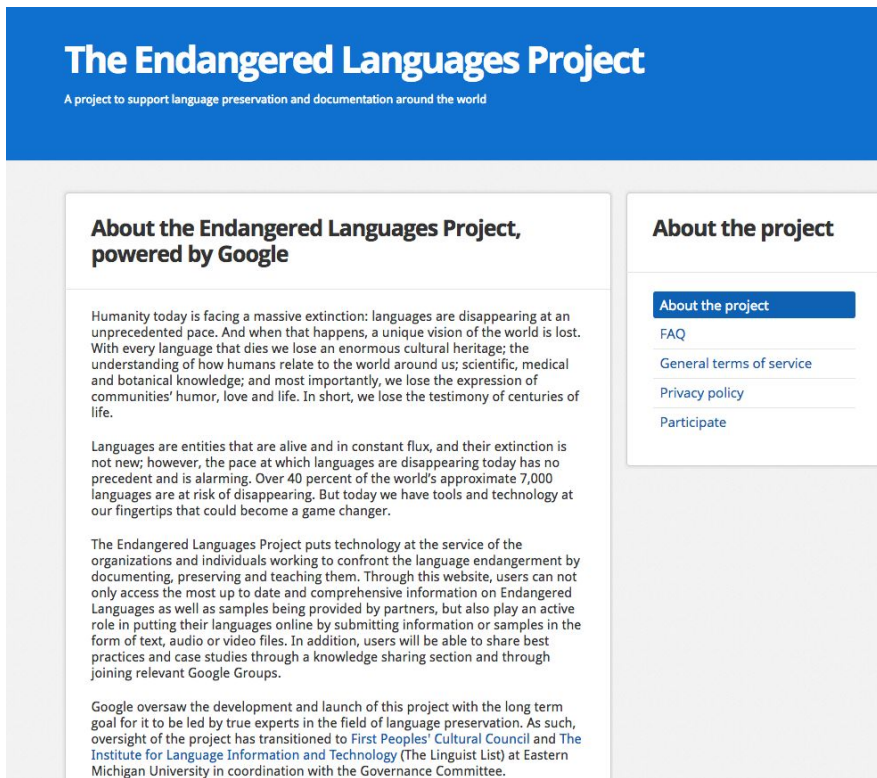


Afrikaans	English	Igbo	Maltese	Somali
Albanian	Esperanto	Indonesian	Maori	Spanish
Arabic	Estonian	Irish	Marathi	Sundanese
Armenian	Filipino	Italian	Mongolian	Swahili
Azerbaijani	Finnish	Japanese	Myanmar (Burmese)	Swedish
Basque	French	Javanese	Nepali	Tajik
Belarusian	Galician	Kannada	Norwegian	Tamil
Bengali	Georgian	Kazakh	Persian	Telugu
Bosnian	German	Khmer	Polish	Thai
Bulgarian	Greek	Korean	Portuguese	Turkish
Catalan	Gujarati	Lao	Punjabi	Ukrainian
Cebuano	Haitian Creole	Latin	Romanian	Urdu
Chichewa	Hausa	Latvian	Russian	Uzbek
Chinese	Hebrew	Lithuanian	Serbian	Vietnamese
Croatian	Hindi	Macedonian	Sesotho	Welsh
Czech	Hmong	Malagasy	Sinhala	Yiddish
Danish	Hungarian	Malay	Slovak	Yoruba
Dutch	Icelandic	Malayalam	Slovenian	Zulu

Endangered Languages Project

Proprietary + Confidential

Google-launched initiative to promote and document info on endangered languages (2012)



The screenshot shows the homepage of 'The Endangered Languages Project'. The header is a solid blue bar with the project name in white. Below the header, the main content area is light gray. On the left, there's a large text block titled 'About the Endangered Languages Project, powered by Google' which contains three paragraphs of text. On the right, there's a sidebar with a blue button 'About the project' and four links: 'FAQ', 'General terms of service', 'Privacy policy', and 'Participate'.

The Endangered Languages Project

A project to support language preservation and documentation around the world

About the Endangered Languages Project, powered by Google

Humanity today is facing a massive extinction: languages are disappearing at an unprecedented pace. And when that happens, a unique vision of the world is lost. With every language that dies we lose an enormous cultural heritage; the understanding of how humans relate to the world around us; scientific, medical and botanical knowledge; and most importantly, we lose the expression of communities' humor, love and life. In short, we lose the testimony of centuries of life.

Languages are entities that are alive and in constant flux, and their extinction is not new; however, the pace at which languages are disappearing today has no precedent and is alarming. Over 40 percent of the world's approximate 7,000 languages are at risk of disappearing. But today we have tools and technology at our fingertips that could become a game changer.

The Endangered Languages Project puts technology at the service of the organizations and individuals working to confront the language endangerment by documenting, preserving and teaching them. Through this website, users can not only access the most up to date and comprehensive information on Endangered Languages as well as samples being provided by partners, but also play an active role in putting their languages online by submitting information or samples in the form of text, audio or video files. In addition, users will be able to share best practices and case studies through a knowledge sharing section and through joining relevant Google Groups.

Google oversaw the development and launch of this project with the long term goal for it to be led by true experts in the field of language preservation. As such, oversight of the project has transitioned to [First Peoples' Cultural Council](#) and [The Institute for Language Information and Technology \(The Linguist List\)](#) at Eastern Michigan University in coordination with the Governance Committee.

About the project

- About the project
- FAQ
- General terms of service
- Privacy policy
- Participate

Language communities & technology

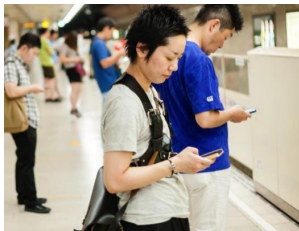
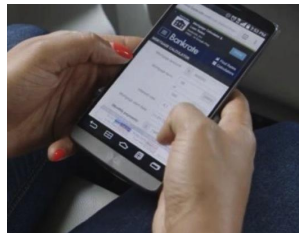


User-driven, needs local engagement

Proprietary + Confidential

Technology is **egalitarian**

Features and launches are **user-driven**




Opportunity & **responsibility** to engage with & use technology available













Open-source initiatives: Common Locale Data Repository

Proprietary + Confidential

Survey Tool 29 Read-Only ⚙ Coverage: Modern Instructions ▾

> Irish / Dashboard

▼ English Changed (38/38) 

Code	English	CLDR 28	Winning 29	Action
Timezones - Europe : Cities and Regions				
Uzhgorod ↻	Uzhhorod OLD: Uzhgorod	Uzhgorod	=	  
Locale_Display_Names - T_Africa : Western Africa				
SH ↻	St. Helena OLD: Saint Helena	San Héilín	=	  
Locale_Display_Names - T_Oceania : Outlying Oceania				
GS ↻	South Georgia & South Sandwich Islands OLD: South Georgia and the South Sandwich Islands	An tSeoirsia Theas agus Oileáin Sandwich Theas	=	  
HM ↻	Heard & McDonald Islands OLD: Heard Island and McDonald Islands	Oileán Heard agus Oileáin McDonald	=	  
Locale_Display_Names - Languages_E_J : E				

CLDR is a Unicode project used widely across the web: supported and utilised by Google, Microsoft, IBM, Apple etc.
Open to all, contributions welcomed by language bodies

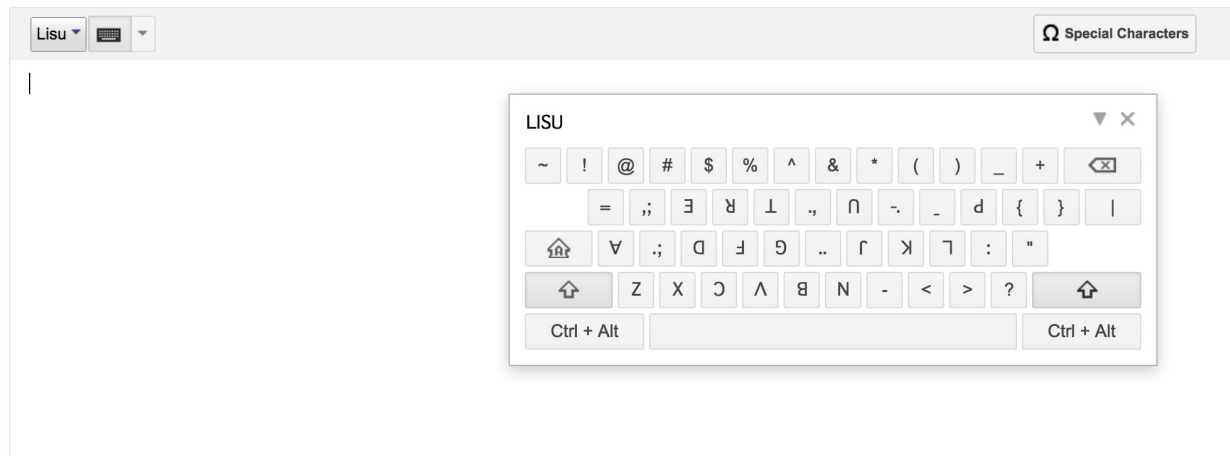
Technology as an enabler

e.g. **Lisu** keyboard as Google intern project

Try Google Input Tools online

Google Input Tools makes it easy to type in the language you choose, anywhere on the web. [Learn more](#)

To try it out, choose your language and input tool below and begin typing.



Challenge now: connectivity, not the language

Thank you
Go raibh maith agaibh