

# Machine Translation APIs

Marc Johnson

[Translations.com](https://www.translations.com)

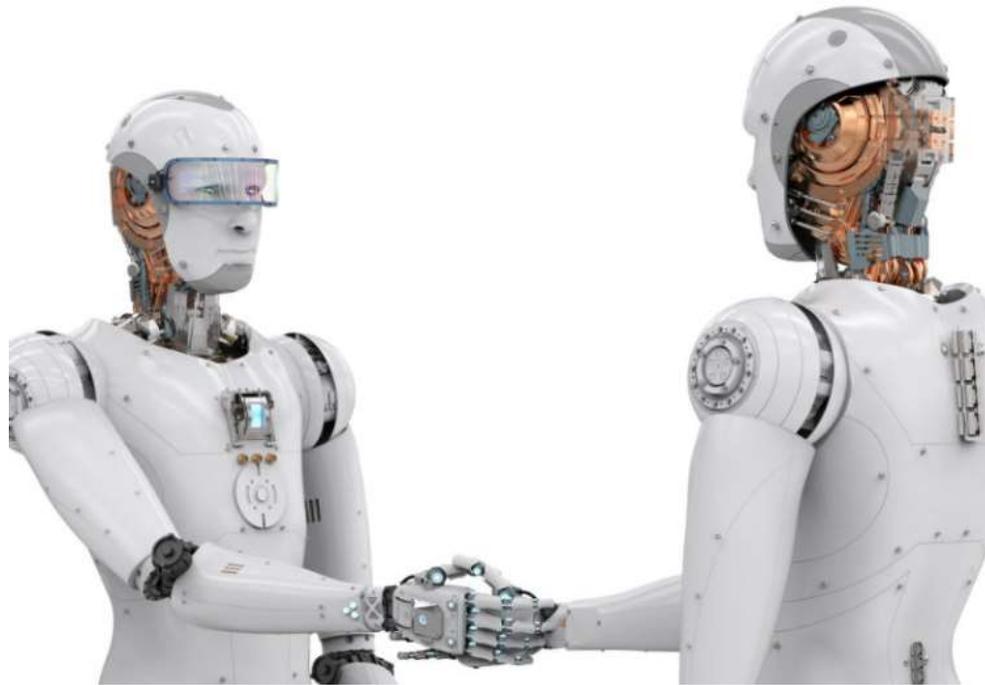
# Who is this guy?



 **translations.com**

 **oneLink**

# APIs



# Application Programming Interfaces

- Rules
- Build your application then follow the rules
- Lots of flavors
  - SOAP
  - RPC
  - REST

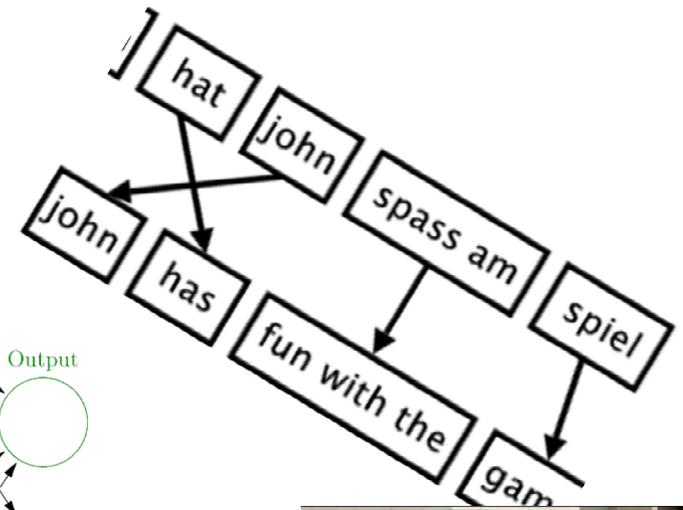
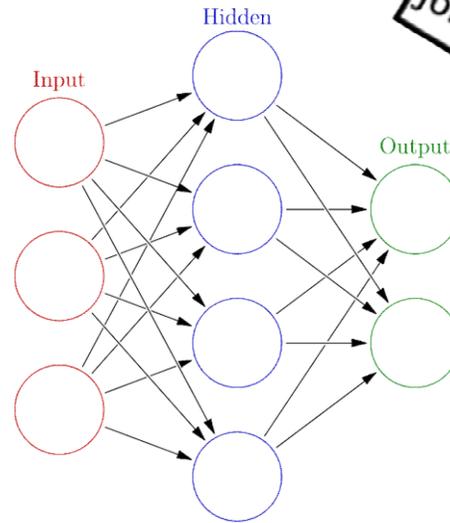
# REST APIs



# Machine Translation

The sky is blue = Le ciel est bleu

$$\sum_i y_i = n \cdot a + b \cdot (\sum x) + c \cdot \sum_i (x_i)^2$$



# Example Based MT

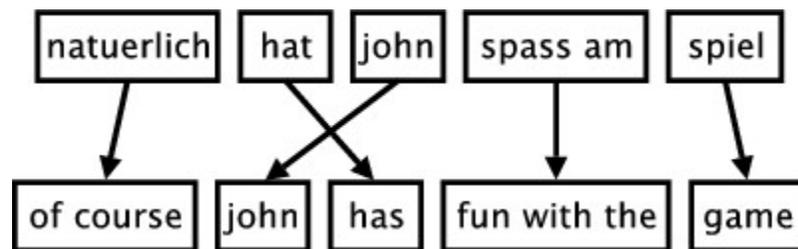
- Translation Memory tools
- 100% match and can measure the percent difference:
  - The sky is blue = Le ciel est bleu
  - The sky is blue = 100% match
  - The sky is **grey** = 75% match

# Cons

- It is green
- Il est vert
- Elle est vert

# Rules Based MT

- Dictionary of words, can be a specific domain.
- Mapping rules for the source and target languages



# Cons

- Languages with very different syntax don't map well.
- Can have same issue EB has, but rules help.

# Statistical MT

- Mathematical models built around analysis of language corpuses (or corpora).
- Literally looking at the likelihood of two words appearing next to each other in a sentence.

# Cons

- Lots of processing power.
- Takes time to build model.
- Can only look at a few words for context.
- Bad results if corpora has bad data.

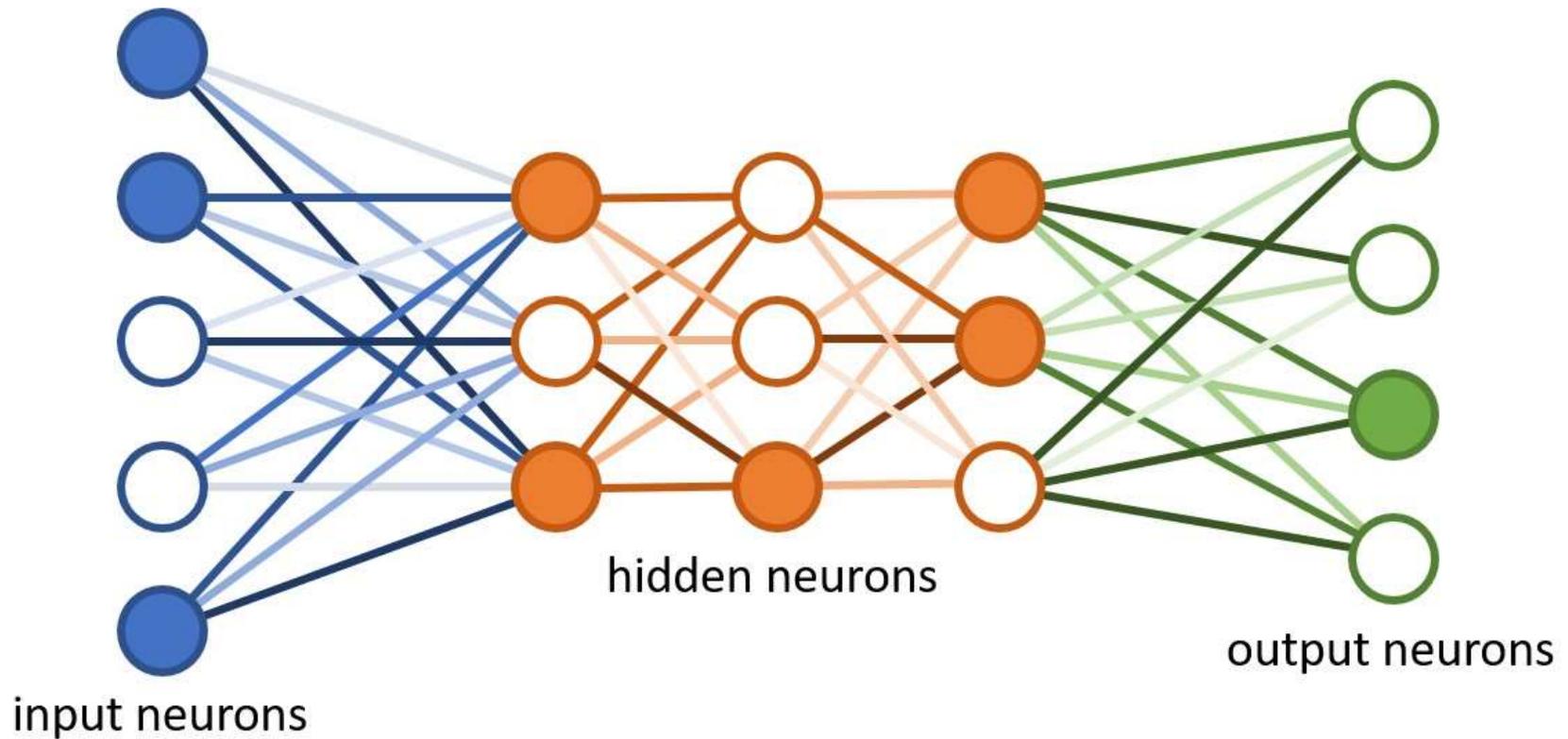
# Hybrid Machine MT

- Combo of RBMT and SMT.
- Often a multi-pass system.
  - RB translation followed by a statistical analysis.

# Cons

- See Slide 12.

# Neural MT



# Cons

- Slow to build neural model.
- “Slow” to translate.

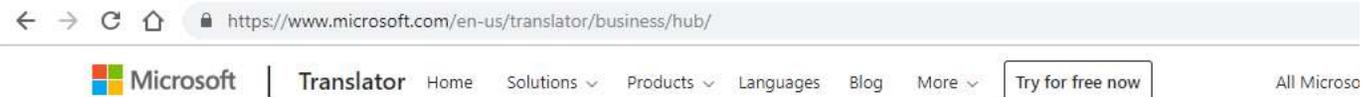
# Microsoft MT

- REST API
- SMT or NMT (you choose)
- You can choose categories for SMT.
- Neural network models available for API by using the 'generalnn' category ID.
- First 2,000,000 chars free per month.

# Training MT with Microsoft

- You can train SMT models with your own corpora.
- You call the model by using special API key.

# Hub



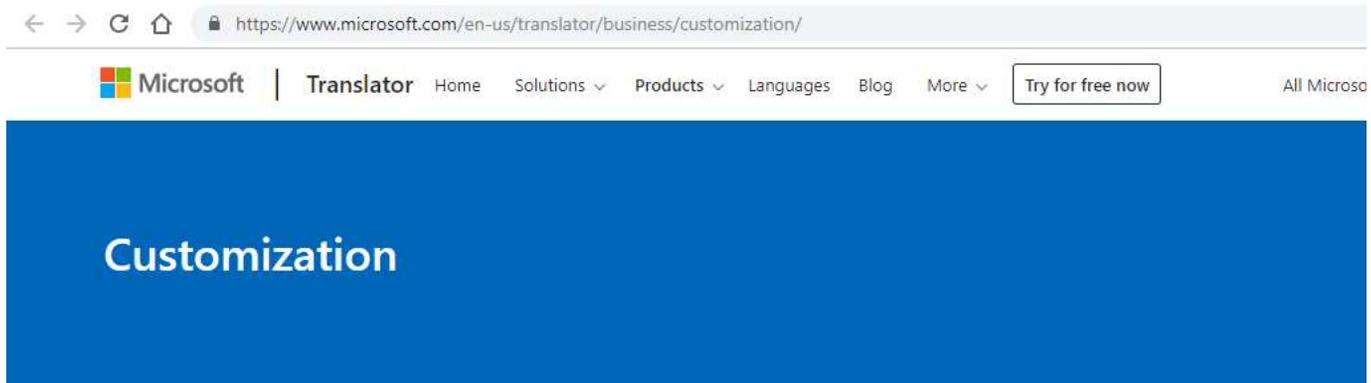
## Microsoft Translator Hub

Microsoft Translator Hub will be retired on April 30, 2019. Please review the following information to ensure your Hub projects are properly migrated to Custom Translator.

The Microsoft Translator Hub, an extension of the Microsoft Translator service, is used to create customized statistic machine translation systems for use with the Translator Text API V2.

The Translator Text API V3 was [released](#) in May 2018, offering customers Microsoft Translator's advanced neural machine translation. [Custom Translator](#) is used to create customized neural machine translation

# Custom Translator



Custom Translator is a feature of the Microsoft Translator service. With Custom Translator, enterprises, app developers, and language service providers can build neural translation systems that understand the terminology used in their own business and industry. The customized translation system will then seamlessly integrate into existing applications, workflows and websites.

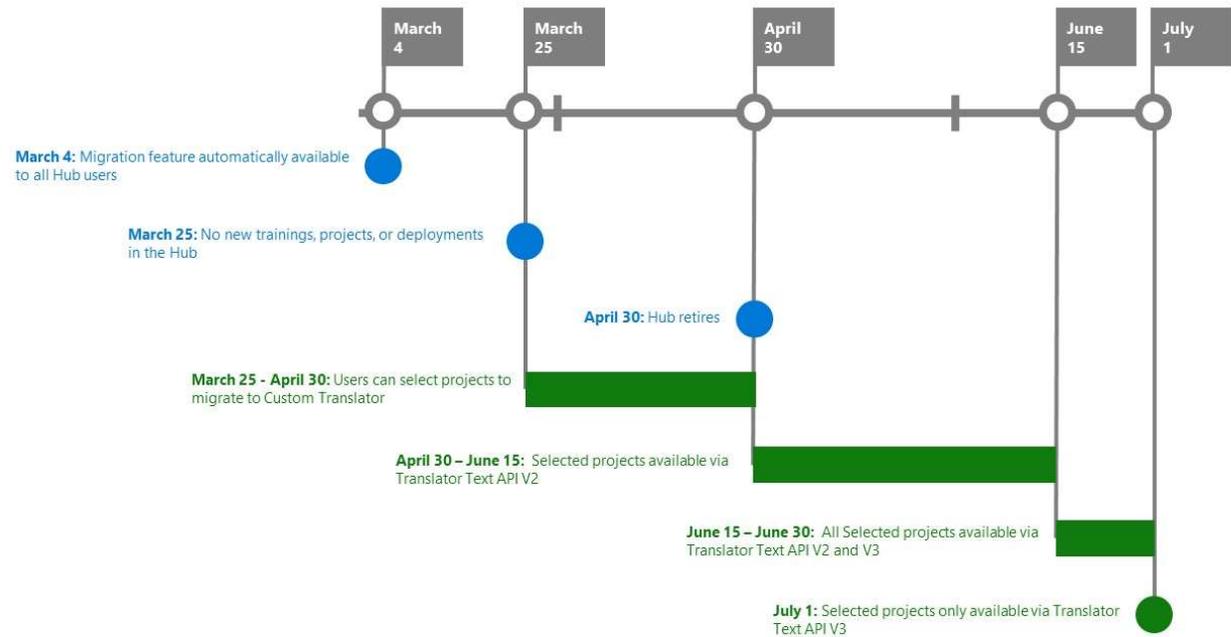
Custom Translator allows users to customize Microsoft Translator's advanced [neural machine translation](#) for Translator's [supported neural translation languages](#). Custom Translator can be used for customizing text when using the Microsoft Translator Text API , and speech translation using the Microsoft Speech services.



# Migration Timeline

Timeline for all language pairs

Timeline for language pairs that are not available in Custom Translator



# Costs

MS	Fee/month	1M Chars of translation:	1M Chars of Custom Translation:	Training: 1M Source +Target Chars	Monthly fee for custom model
S1	\$0	\$10	\$40	\$10	\$10/month/model
S2	\$2,056	\$8.22			
S3	\$6,000	\$6			
S4	\$45,000	\$4.50			
C2	\$2,055	\$10	\$32.88	\$8.22*	
C3	\$6,000		\$24	\$6*	
C4	\$45,001		\$18	\$4.50*	
				*Max \$300/training	

# Google MT

- REST API
- NMT
- \$20 per 1,000,000 characters.
- You can build your own NMT models but it's not cheap.
- AutoML training is \$3 per hour.

# Translations.com MT

- Choice of MT engine:
  - SMT
  - NMT
    - Google
    - Microsoft
    - Custom



# OneLink API Sample Syntax

```
curl -k
  --header 'Host:es-otx.onelink-translations.com'
  --request POST 'https://es-otx.onelink-translations.com/OneLinkOTX/'
  --data
  'otx_mimetype=text/html&otx_account=otx,otxpass&otx_service=t
x&otx_content=<p>I am going for a long walk</p>'
; echo
```

# Uses

- Real time translations
- Agile Workflows
- Testing source material for MT quality

# Real-time translation

Translate

From:

Italian

To:

English

View:

Translation

Original



YAHOO!  
FINANZA

Cerca notizie, codici o aziende.

Search for

 Mail

Home Finance

My wallet

Market overview

Quotes

Personal finance

Industry

Currency converter

Yahoo Original

Cerca quotazioni

## Navigator, Regions in alarm

It is a fragile chain because the rings that compose it have been put  
**to know more »**

Personal portfolio and marke...

[Customize](#)

Recent views >

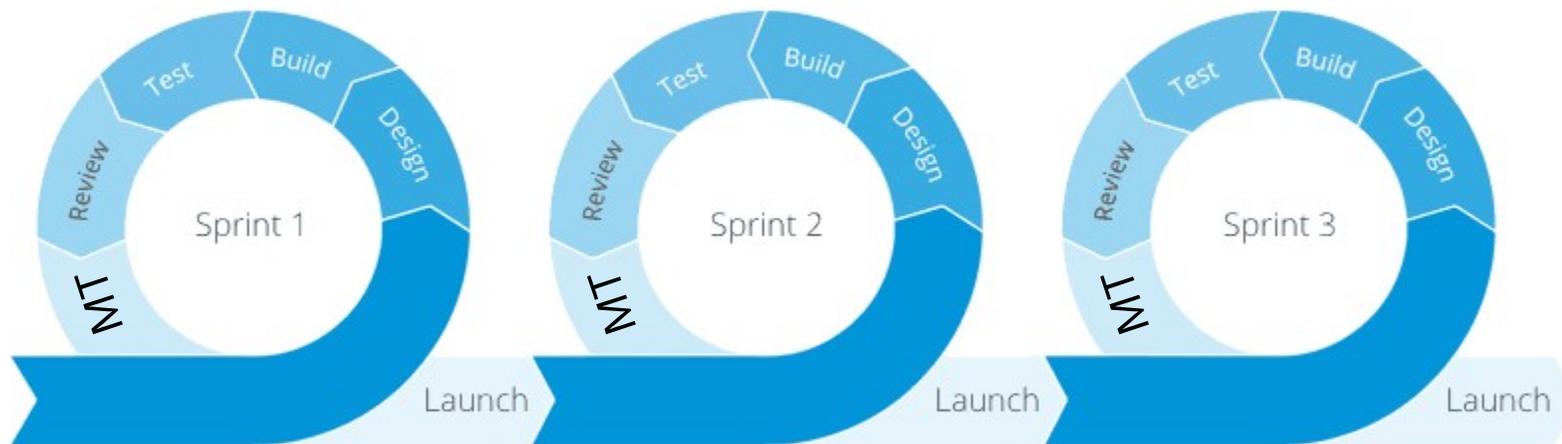
Your list is empty.



# ¿Como?

<p>Not secure   translation2.paralink.com/English-Spa</p> <p>BX Administrati OneLink Billboard OneLink Portal</p> <p>Detected: English</p> <p>English Spanish</p> <p>What is your name?</p> <p>Translate Compare à è i</p> <p>PROMT-Online translation Google translation Micro</p> <p>¿Cómo se llama usted?</p> <p><input type="checkbox"/> Back translation</p>	<p>Not secure   translation2.paralink.com/English-Spa</p> <p>BX Administrati OneLink Billboard OneLink Portal</p> <p>Detected: English</p> <p>English Spanish</p> <p>What is your name?</p> <p>Translate Compare à è i</p> <p>PROMT-Online translation Google translation Micro</p> <p>¿Cuál es su nombre?</p> <p><input type="checkbox"/> Back translation</p>	<p>Not secure   translation2.paralink.com/English-Spanish-Translation/</p> <p>BX Administrati OneLink Billboard OneLink Portal User bug count -</p> <p>Detected: English</p> <p>English Spanish</p> <p>What is your name?</p> <p>Translate Compare à è i ò ù æ</p> <p>PROMT-Online translation Google translation Microsoft translation</p> <p>¿Cómo te llamas?</p> <p><input type="checkbox"/> Back translation</p>
---	---	--

# Agile workflows



# Testing your content for MT

1

## Welcome to Android

### **About Android 5.0, Lollipop**

Android 5.0, Lollipop is the latest version of Android, the operating system that powers not just phones and tablets, but also wearables, TVs, and even cars.

This guide introduces the Android 5.0 software available on Nexus and Google Play edition mobile devices. For hardware details for these devices, such as how to insert a SIM card, specs, and so on, visit [support.google.com/android](http://support.google.com/android). There, you can find comprehensive information about all types of Android phones and tablets, including more details about Nexus and Google Play edition devices and other versions of Android.

# Creating Content for MT

- Keep sentences short.
- Run spell check.
- Keep terminology consistent.
- Don't emphasize with ALL CAPS.
- Avoid & and / to get your meaning across.
- Don't be funny, trite, or use slang.

# Questions?

