



Use More Machine Translation and Keep Your Customers Happy

A Practitioner's Perspective

Glen Poor, Principal Program Manager
Global Service eXperience team, Office Product Group





Office



Word



Excel



PowerPoint



Outlook



OneNote



Publisher



SharePoint



OneDrive
for
Business



Skype
for
Business

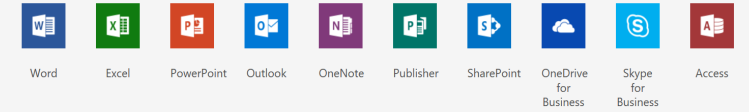


Access



Office on PCs, tablets, and phones

Office



Office on PCs, tablets, and phones



Online Meetings



Advanced email



Intranet and team sites



Intelligent search and discovery



Meeting Broadcast



Document and email access control



Corporate social network



Enterprise video service



Audio Conferencing



Advanced information protection



Office Online



Voicemail integration (Unified Messaging)



Modern voice with Phone System



Threat intelligence



Work management



Intelligent compliance solutions



Instant messaging and Skype connectivity



Advanced security



Professional digital storytelling



Apps for Office and SharePoint



Hub for teamwork



Analytics tools



Mobility



Workday management



Email and calendars



File storage and sharing



Enterprise management of apps



Workflow automation



Office

50% +

of Office Revenue \$ is
International

90+

Languages

200+

Markets

50% +

of O365 Commercial Usage is
International

1,000,000,000+

Words localized each year

Why so many languages?

Revenue is in 40 languages

Enterprise, Education, Small Business and Consumer

High usage of English and other dominant languages in small markets

Accessibility & Citizenship

Governments want to support minority languages

Respect people who want to support their culture

Language contributes to Digital Divide

Small market may be lots of speakers who don't use our products, yet

MT could help more...

Hard sell to Finance when cost per word for a non-revenue market >> revenue market

Low Resource language research AND cooperation

90+

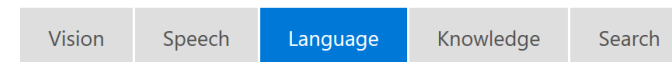
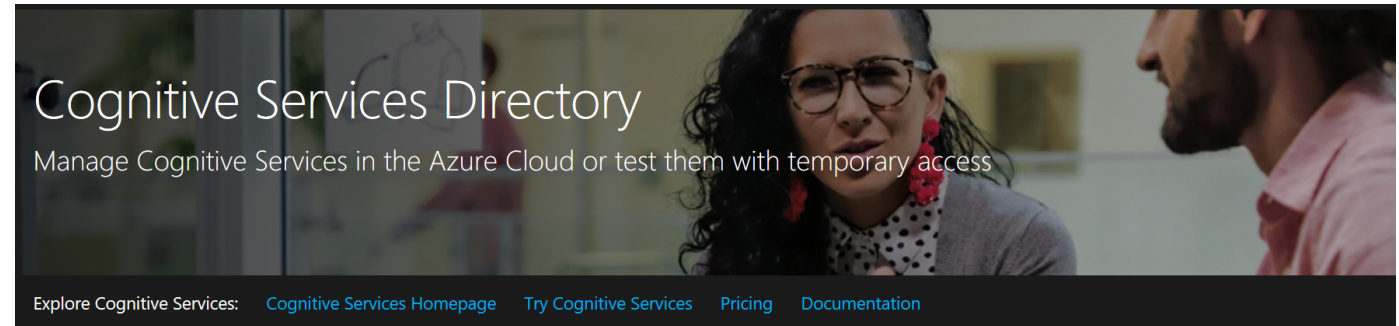
Languages

200+

Markets

We use Microsoft MT

We use the Cognitive
Service APIs for Translation
features in the product



Language Understanding (LUIS)

Teach your apps to understand commands from your users

[Try Language Understanding \(LUIS\)](#) | [Use with an Azure subscription](#)



Bing Spell Check API

Detect and correct spelling mistakes in your app

[Try Bing Spell Check API](#) | [Use with an Azure subscription](#)



Web Language Model API **PREVIEW**

Use the power of predictive language models trained on web-scale data

[Try Web Language Model API](#) | [Use with an Azure subscription](#)



Text Analytics API

Easily evaluate sentiment and topics to understand what users want

[Try Text Analytics API](#) | [Use with an Azure subscription](#)



Translator Text API

Easily conduct machine translation with a simple REST API call

[Use with an Azure subscription](#)



Linguistic Analysis API **PREVIEW**

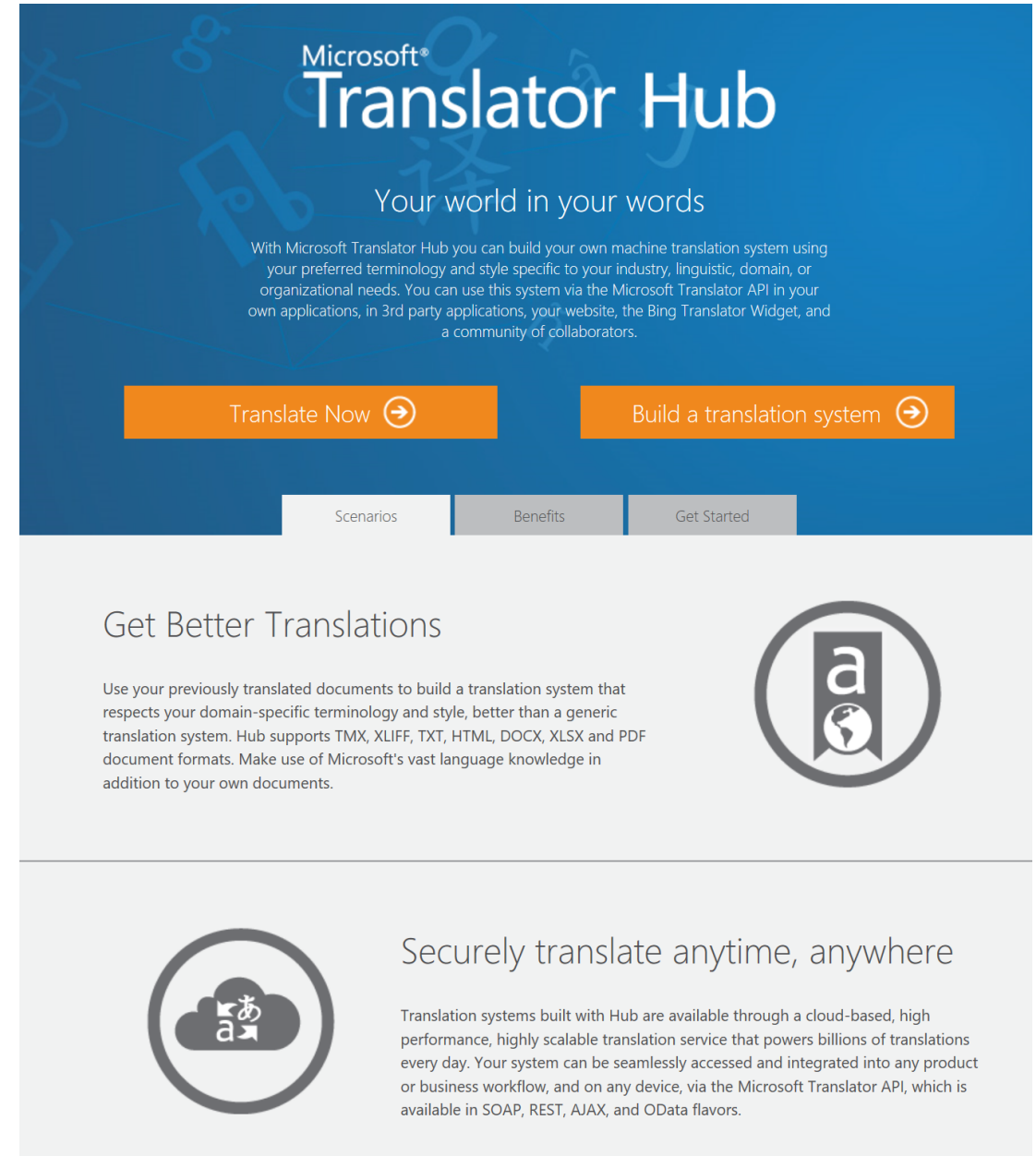
Simplify complex language concepts and parse text with the Linguistic Analysis API

[Try Linguistic Analysis API](#)

We use Microsoft MT

MS Translator supports 60
languages

We use domain tuned
engines for 38 languages
and are adding more.



Microsoft
Translator Hub

Your world in your words

With Microsoft Translator Hub you can build your own machine translation system using your preferred terminology and style specific to your industry, linguistic, domain, or organizational needs. You can use this system via the Microsoft Translator API in your own applications, in 3rd party applications, your website, the Bing Translator Widget, and a community of collaborators.


Translate Now →

Build a translation system →

Scenarios Benefits Get Started


Get Better Translations

Use your previously translated documents to build a translation system that respects your domain-specific terminology and style, better than a generic translation system. Hub supports TMX, XLIFF, TXT, HTML, DOCX, XLSX and PDF document formats. Make use of Microsoft's vast language knowledge in addition to your own documents.



Securely translate anytime, anywhere

Translation systems built with Hub are available through a cloud-based, high performance, highly scalable translation service that powers billions of translations every day. Your system can be seamlessly accessed and integrated into any product or business workflow, and on any device, via the Microsoft Translator API, which is available in SOAP, REST, AJAX, and OData flavors.



Office as a Service Changes Everything



3 Years → Constant Release

Packaged Product → Service

Hard to update → Frequent Updates

Be Perfect → Good Enough and React

BUFT → Synthetic + Analytic Validation

Occasional → Constant Customer Contact

Pre Release Usability → Data Driven and A/B Studies

Constant focus on Customer Satisfaction



Localization as a Service

A resource enters, gets translated, validated and published without human hands touching it.



October 28 - November 1, 2016

MT Thresholding: Achieving a defined quality bar with a mix of human and machine translation

Dag Schmidtke

Senior Program Manager, Office Global Services & Experiences
Microsoft Ireland



Our experience of applying MT on Support.Office.Com

Help and support articles

The Assertion

We care about Customer Satisfaction (CSAT)

“Was this information helpful? Yes / No”

We increase MT use so long as CSAT close to HT CSAT

We control the impact of MT with min threshold of recycled text in an article and choosing non-critical articles.

We set the threshold based on reviewing aggregate CSAT

Lower thresholds mean we MT more text

We start conservatively and iterate over time to increase threshold



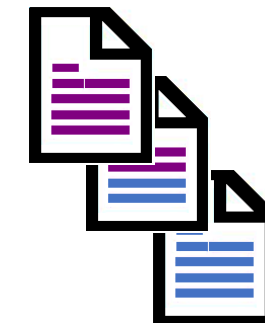
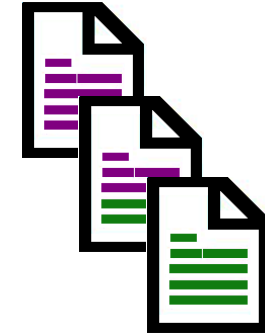
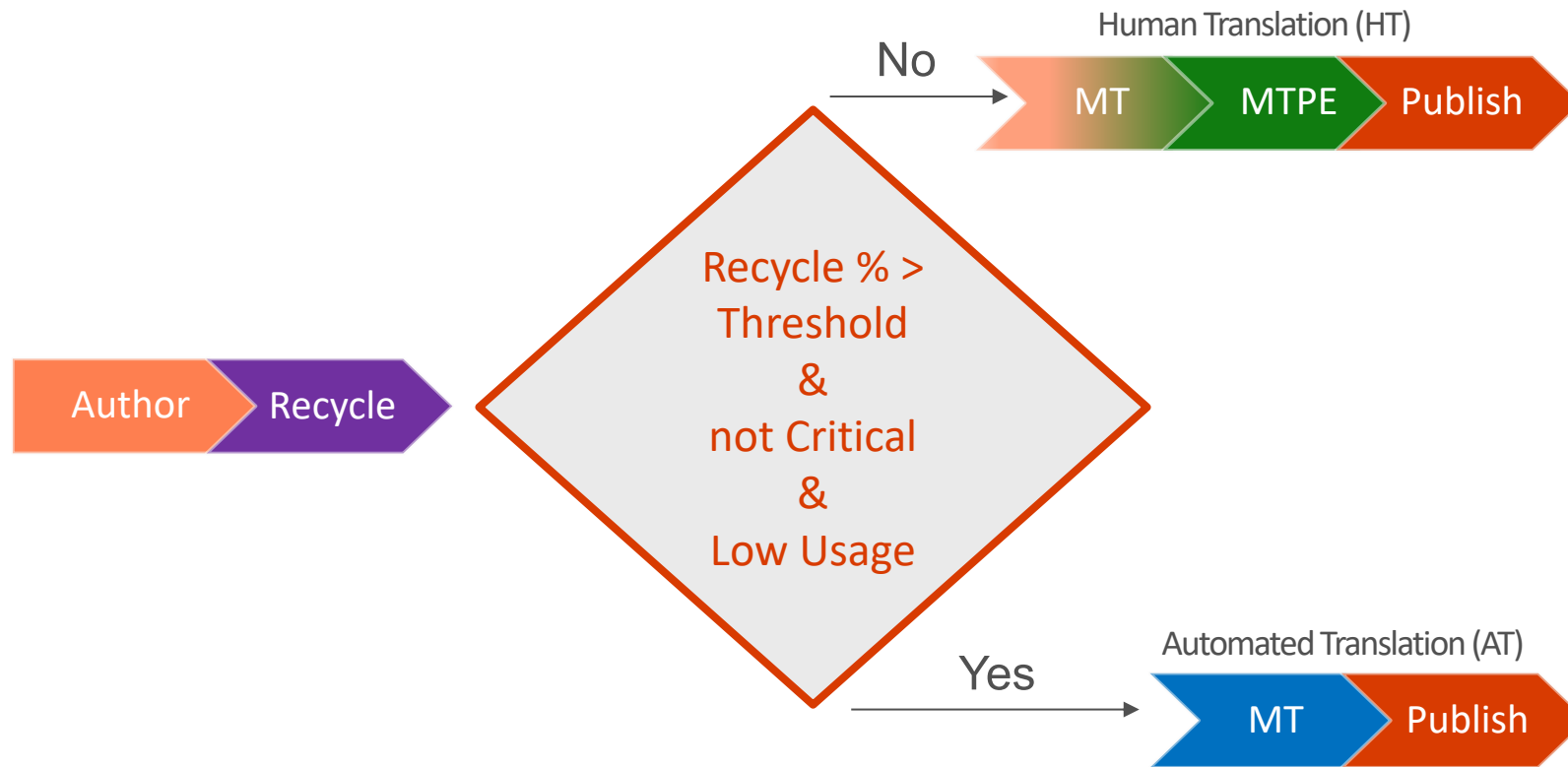
MT Thresholding: Achieving a defined quality bar with a mix of human and machine translation

Dag Schmidtke
Senior Program Manager, Office Global Services & Experiences
Microsoft Ireland



A quick simplified* review

Data Driven Decision

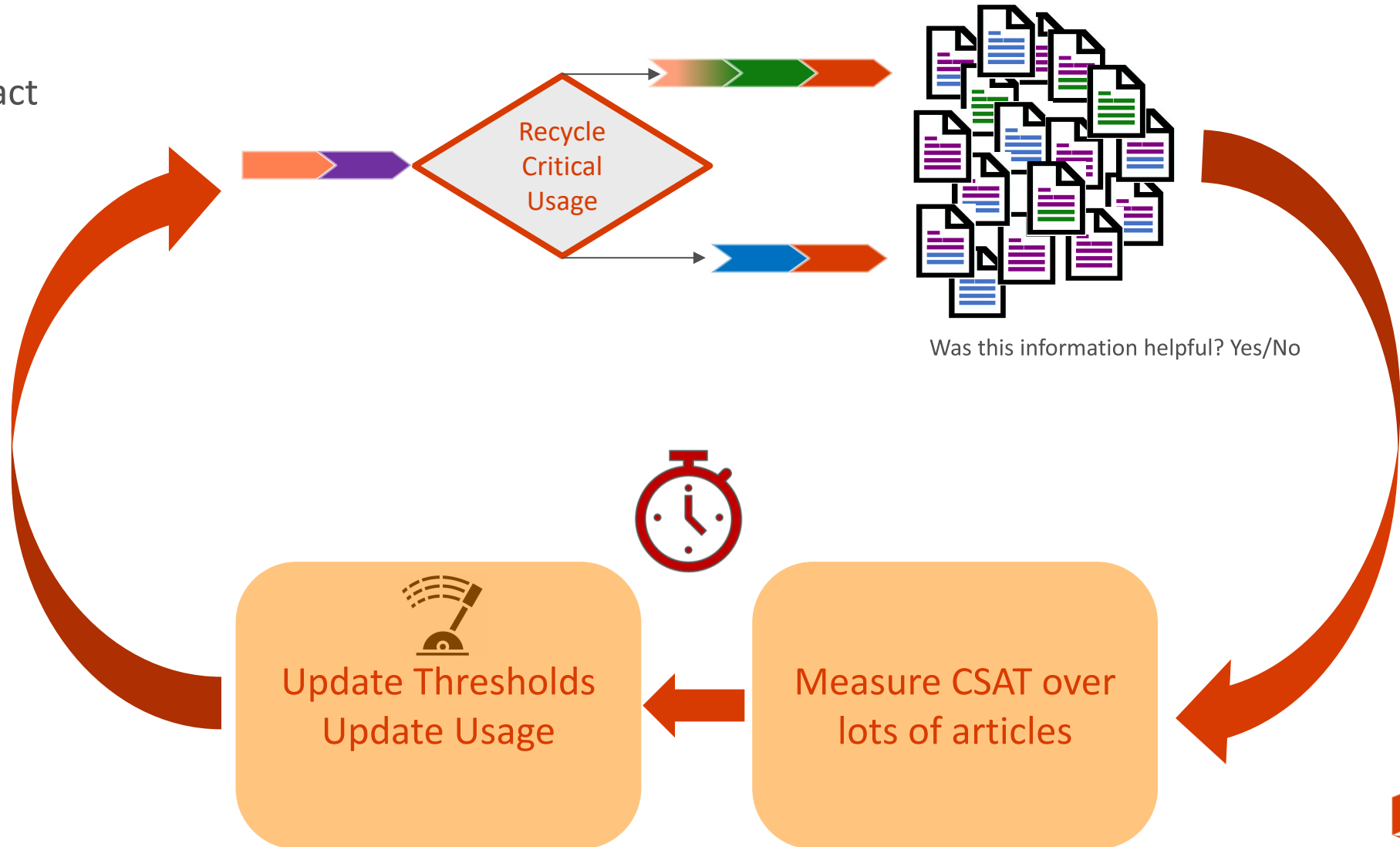


*The decision logic has more wrinkles and stages than shown here.

A quick simplified review

Iterate

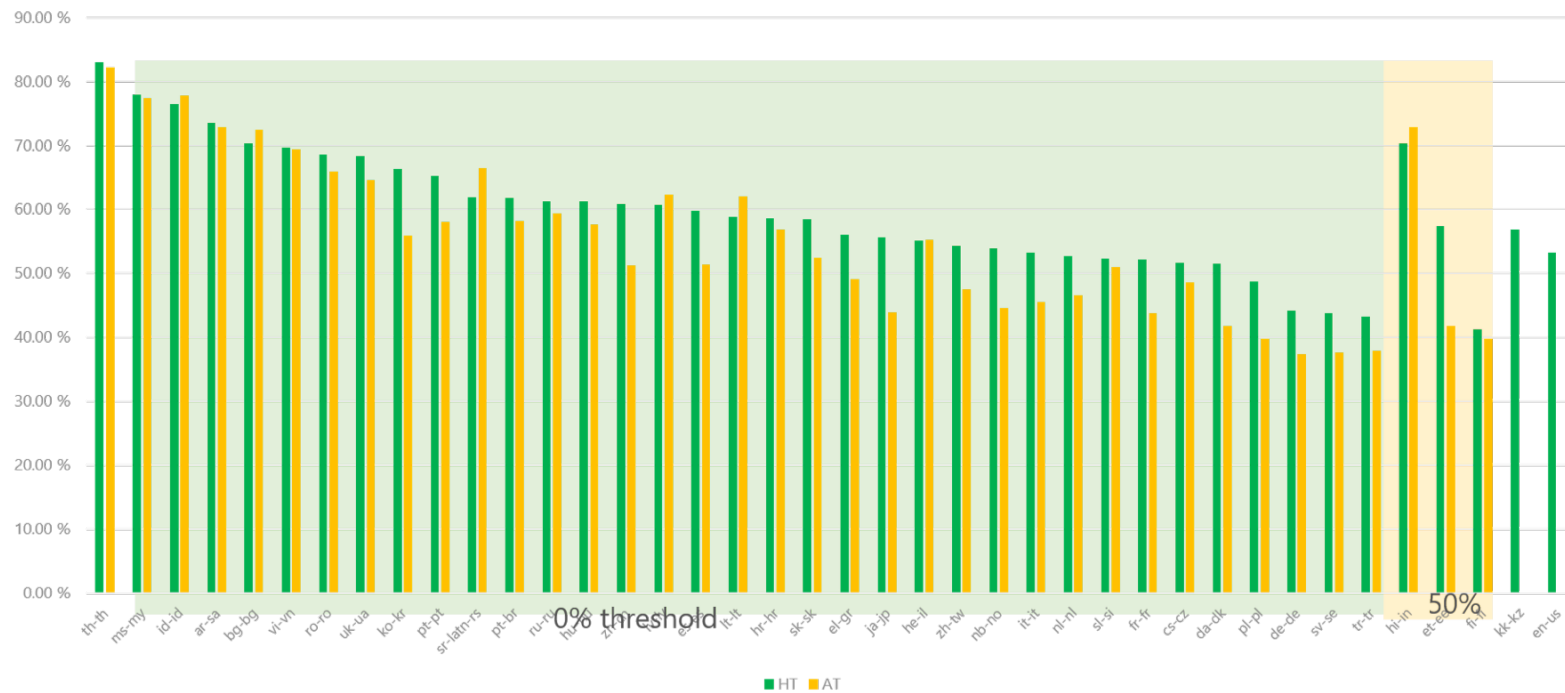
Publish, listen, react



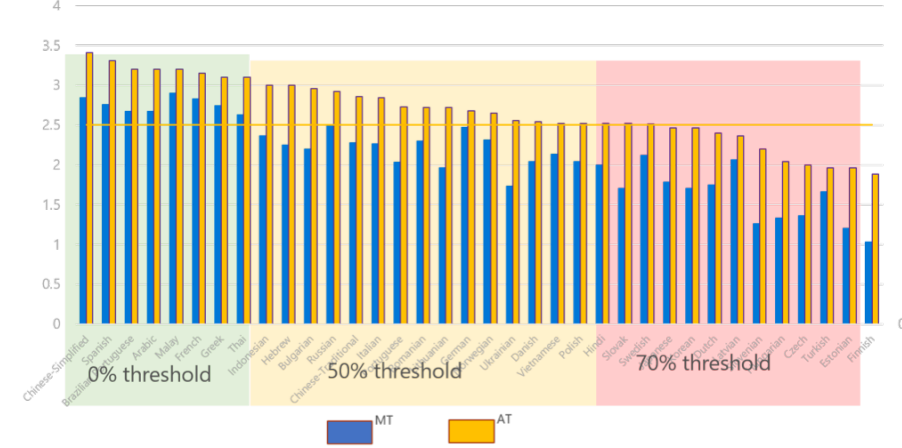
The Journey

2014 - 2017

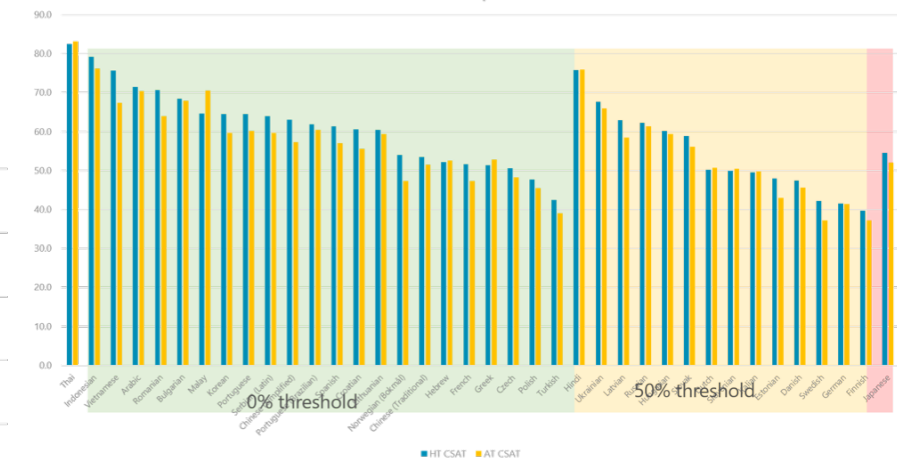
HT v AT CSAT, December 2017



Machine translation quality, SOC evaluation, Oct 2014



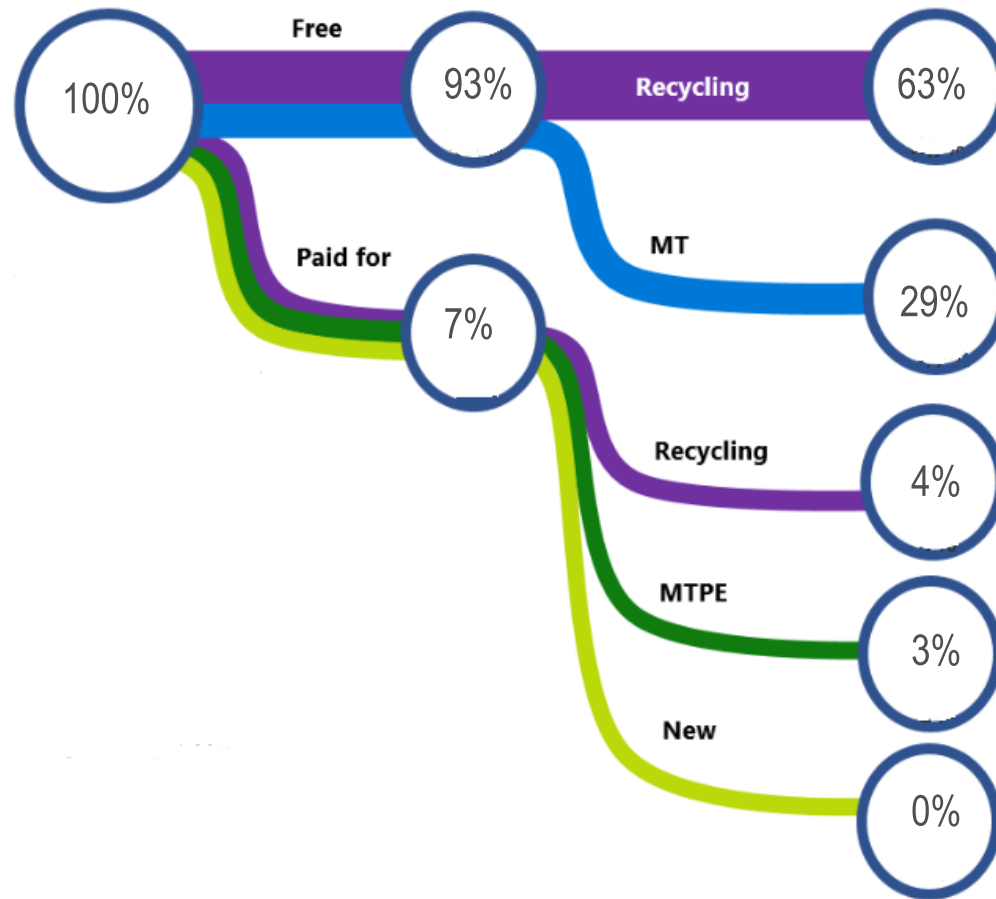
HT v AT CSAT, April 2017



Flow Stats

Word Count by Translation Type

Goal: Maximise use of Recycling and Machine Translation, while protecting Customer Satisfaction



Why does this work?

The MT often works well enough

And when it doesn't, we can hear this and upgrade the article to MTPE

For most articles, we have a lot of recycling

Averaging 54% on AT articles

Customers can get past poor MT if there is enough good translation around it.

Most traffic is focused on a small number of articles

We don't always know what will be important or heavily used, so iterate fast is key

Can we apply to Software?

Software and Content are similar and different

Content is a whole article that we can look at and is a single user experience.

Software is lots of short strings that combined make up an experience we only see at run time.

Both have lots of words that are rarely seen

What does Customer Satisfaction mean for Software?

No equivalent “Was this helpful?” for each string

Net Promoter Score: “Would you recommend this product to a friend or family?”

Net Star Rating: Sentiment rating provided where customers get app from a store

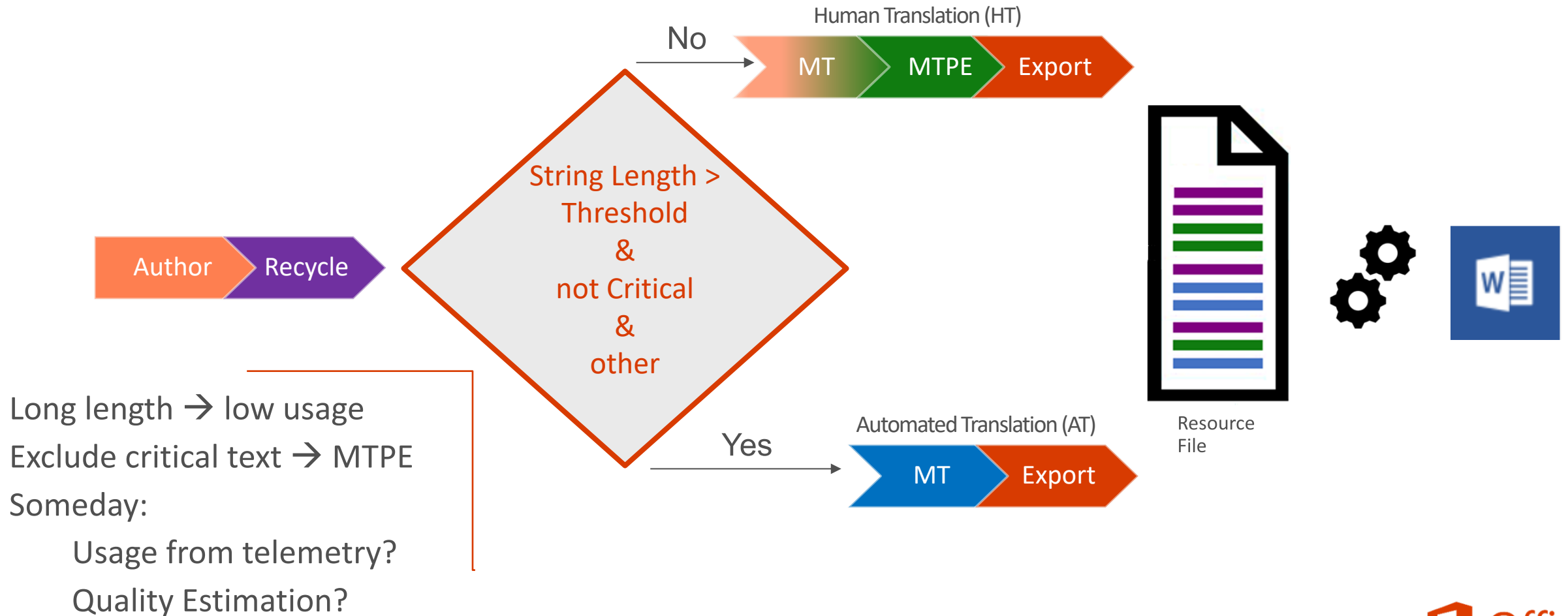
Net Language Quality Score survey: “Are you satisfied with the French language quality?”

Usage Distribution

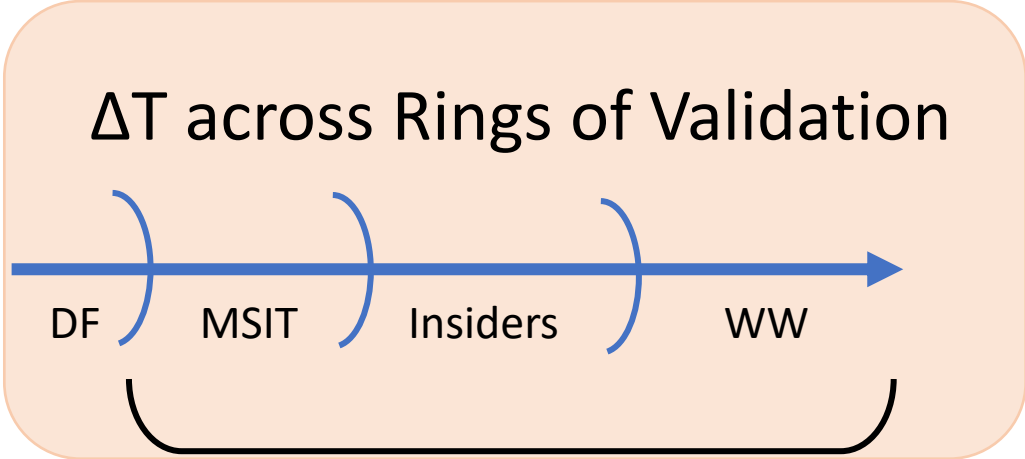
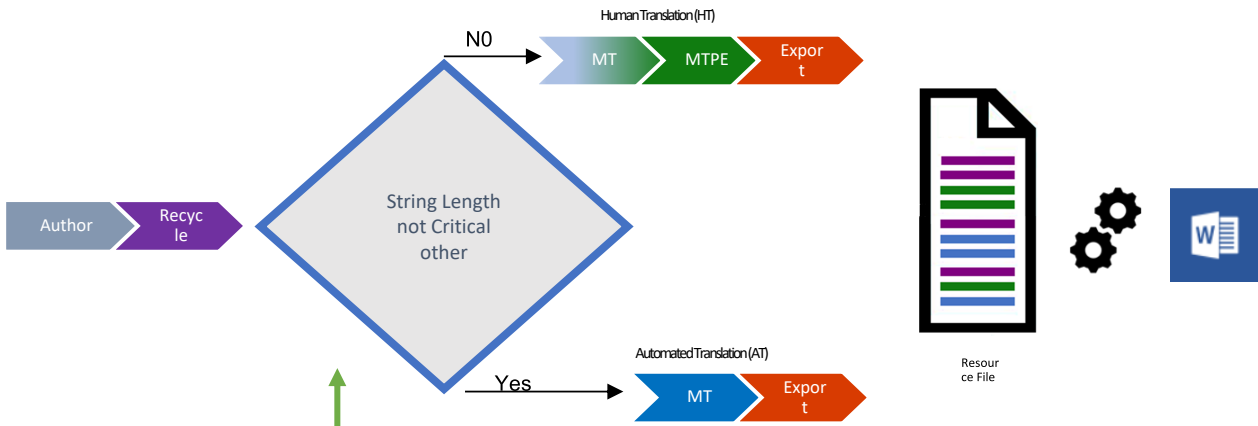
Customer Feedback

A quick simplified* review

Data Driven Decision



*The decision logic has more wrinkles and stages than shown here.



Listen to:
Customer Feedback
NPS & NLQS

Periodic Synthetic
Review

- String Length
- More Languages
- More Products
- Less Exclusion



Happy Customer
Dashboard



The Challenges

Requires substantial changes to systems

We will go live “soon”

Getting nuanced useful Customer Feedback is hard

When we are awful, we hear it quick

Surveys and other feedback channels are mostly not-actionable

Absence of evidence not evidence of absence...

What is “Good Enough” quality bar?

Any customer use of product involves many touch points, what impact does an individual string have?

Office app diversity

No single solution seems to solve for all apps and end points

Customer Listening

Office Customer Voice

Modern software development

Emphasis on getting features to customers, listen and react fast
Feedback and Telemetry impact business decisions

Lots of feedback

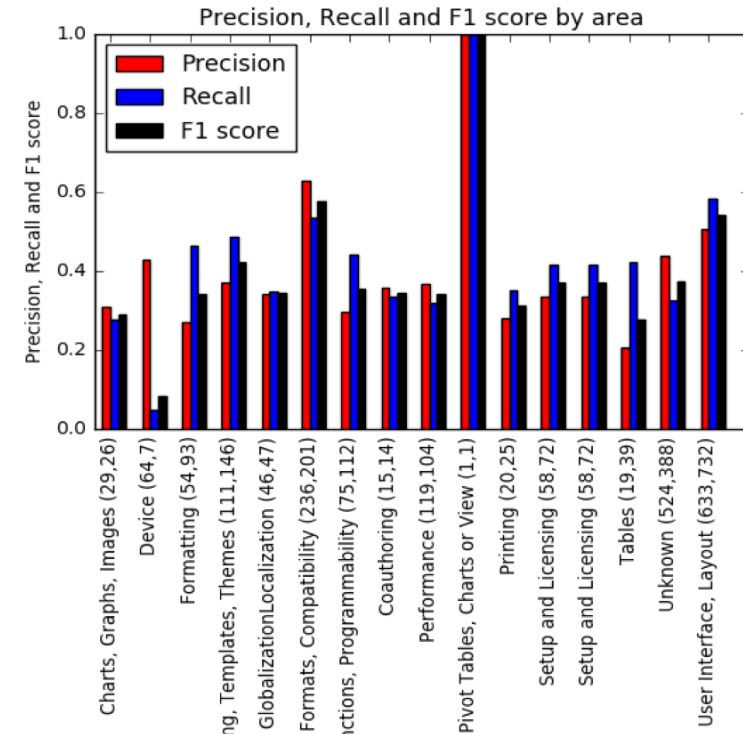
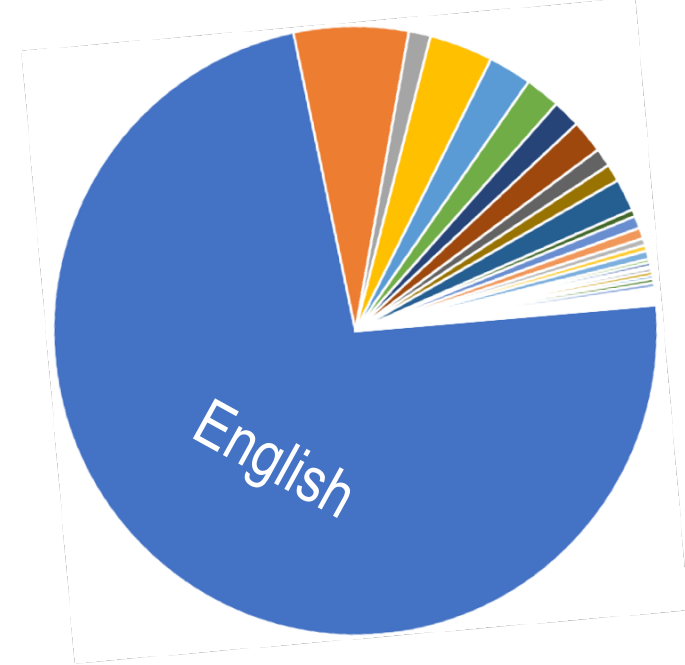
Sources like Forums, Feedback Verbatims, Store Reviews, etc
7.6 million individual pieces past 6 months

Microsoft speaks English

Classifying, monitoring, responding, anomaly detection

Can MT give these customers a voice?

Evaluated influence of different domain tuned MT engines
Neural MT vs domain tuned SMT



Customers Need Help

Virtual Agents are here

In English, for the USA


Intents and Entities


Language Understanding (LUIS)

Proprietary Internal Tooling

Can MT help us?


We will be working to find out.

 Virtual agent



Hello Glen!

I'm Microsoft's Virtual Agent. I'm still in training and I'd love to help you. You can also ask to talk to a person at any time. Please briefly describe your issue below.

e.g.: Reset my Microsoft account password 

Customer Facing

Empower users to work in non-native and across languages seamlessly

“Second-language assistance” – get quick translations while collaborating

+Alt translations, bilingual definitions, read aloud, rich inserts

Translate full email, document, page, or chat

Focus on review, collaboration, and contextual suggestions

Translate live speech & display subtitles in any language

Call/meeting transcript in any language

Presentations across languages

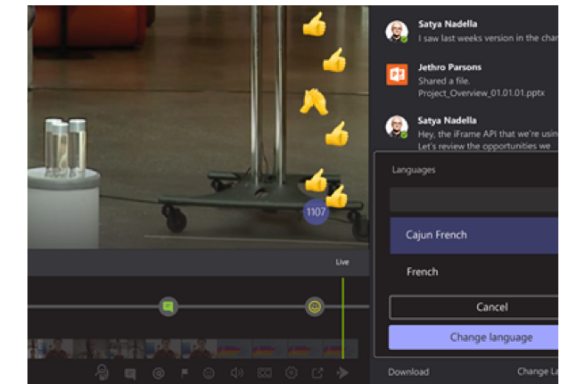
On the phone, the PC, and the web



Contextual email translation – OWA (in-development)



Presentation Translator live captioning – PPT Win32 (Garage add-in)



Transcript feed – Teams Broadcast Office (vision)

Reflections

Will Neural save us?

We have not deployed NMT

We are waiting for domain tuning in Translator Hub

NMT will be better and it will still be variable

We plan to still use Customer Feedback to Calibrate.

We have anecdotal info that Adequacy is more important than Fluency for us.

We will continue to bet on Recycling

We have high recycling percentages and don't see MT replacing those.

We have 90 languages!

Low resource language solutions needed

Will Neural QE help?

We want to make sure that we start with good enough

Thank You

