

Evaluation Report for Customization of Amazon Translate Active Custom Translation with TAUS Financial Domain Data

Language Pair

English → Czech

Domain

Financial

Introduction

Online machine translation engines provide easy access to high quality machine translations. These machine translation engines are optimized for content like news articles and social media posts that end users of these online platforms frequently translate.

Businesses often want to translate text with a different style and a specific topic. For enterprise use, online machine translation engines offer customization with existing translations that reflect the desired style and topic.

TAUS makes such customization data available via the TAUS Data Marketplace and TAUS Matching Data platforms, and now AWS Marketplace.

So that business can clearly assess the value of TAUS data, Polyglot Technology LLC was tasked to independently evaluate the quality of machine translation of Amazon Translate customized with TAUS Data (using [Amazon Translate Active Custom Translation](#)) compared to non-customized Amazon Translate.

BLEU Scores for Amazon Translate and Amazon Translate Active Custom Translation

Machine Translation Evaluation

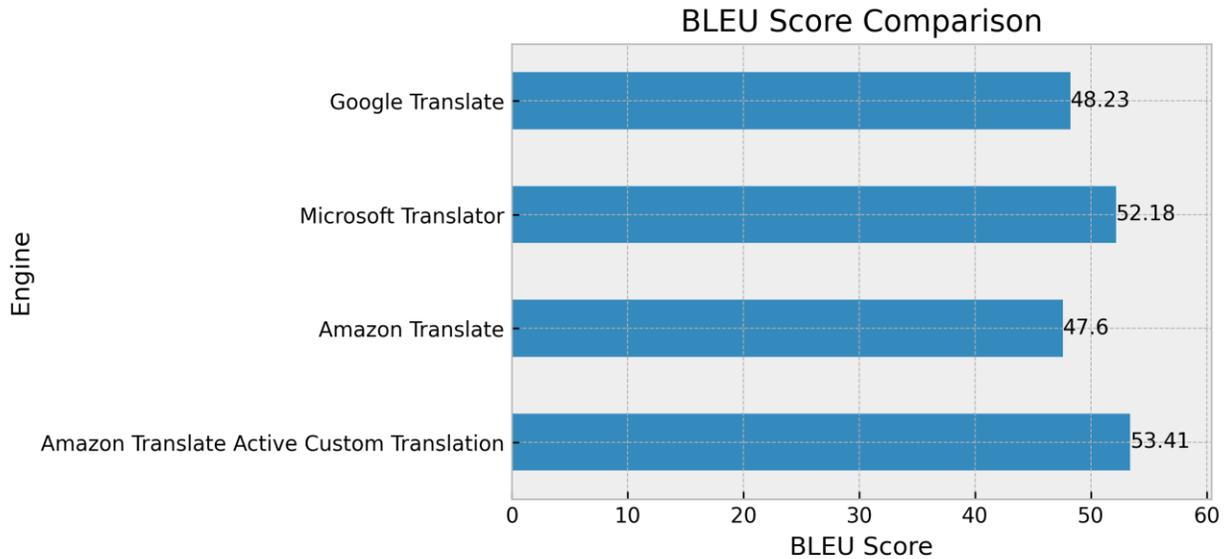
To judge whether machine translation is good or not, human evaluation is the best method. We can ask speakers of the source language and the target language, or better professional translators, to judge whether a machine translation is an adequate and fluent translation of the original text. Or we can ask how close the machine translation is to a human reference translation. Human evaluation however, is slow and hard to scale along language pairs and domains.

Automatic metrics that also use human reference translations have been developed to calculate a numeric score for machine translation quality. For close to 20 years the predominant automatic metric is BLEU, measuring the similarity of machine translations to human reference translations on a scale from 0 to 1 (or 0 to 100 when expressed as percentages). More details on BLEU and how to interpret it can be found in the section “Interpreting BLEU Scores” below.

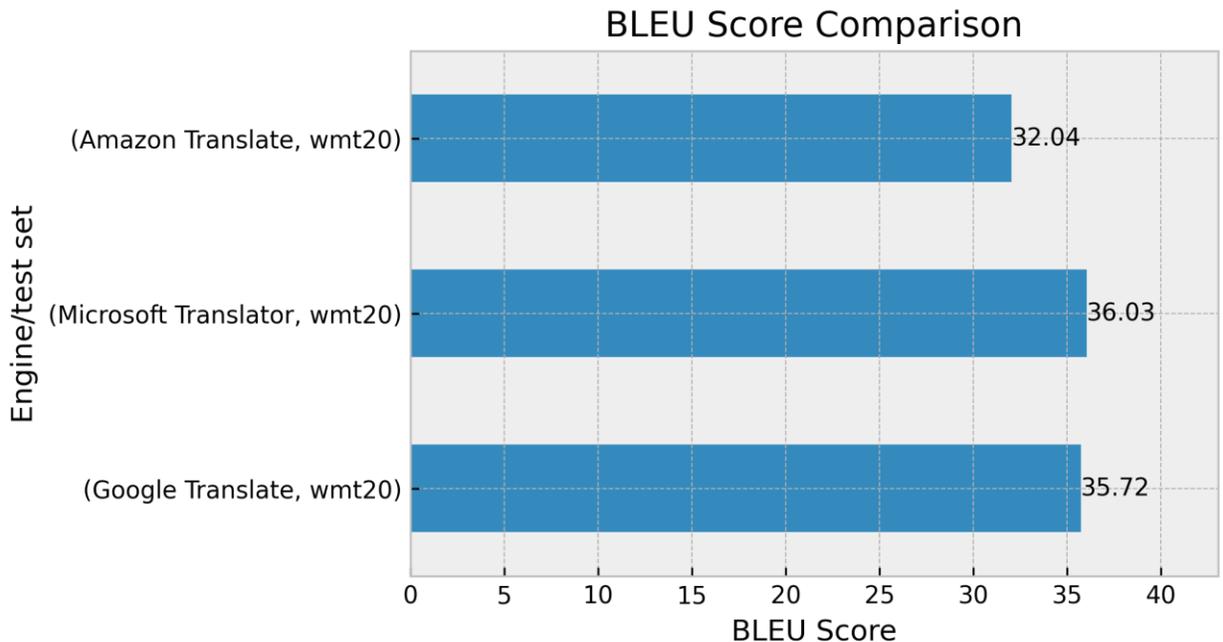
TAUS Test Set

TAUS selects the machine translation customization data by querying its large repository of high-quality translation data with a domain-specific text. The resulting customization dataset is then split at random into a larger training set for Amazon Translate Active Custom Translation and a smaller 2,000 sentence test set that was provided to Polyglot Technology for evaluation.

BLEU Score Results for the TAUS Test Set



BLEU Score Results for Publicly Available Test Sets for the Language Pair



Description of Publicly Available Test Sets

wmt: News text test sets published by the Conference on Machine Translation¹

iwslt: Transcribed TED talk test sets published by the Conference on Spoken Language Translation²

Use Case Specific Evaluation

When employing machine translation for a specific use case, it is advisable to evaluate the systems with usage-scenario specific source text and its human reference translation. Maybe you have already data from a previous, similar project, or your translation vendor can help you create the test data. Polyglot Technology can assist in implementing a robust evaluation program.

Interpreting BLEU Scores

The paragraphs in this section are adapted from Google AutoML Translate's [documentation page on evaluation](#) which is licensed under the [Creative Commons 4.0 Attribution License](#).

[BLEU \(BiLingual Evaluation Understudy\)](#) is a metric for automatically evaluating machine-translated text. The BLEU score is a number between zero and one that measures the similarity of the machine-translated text to a set of high quality reference translations. A value of 0 means that the machine-translated output has no overlap with the reference translation (low quality) while a value of 1 means there is perfect overlap with the reference translations (high quality).

It has been shown that BLEU scores correlate well with human judgment of translation quality. Note that even human translators do not achieve a perfect score of 1.0 (for the reason that a source sentence can have several valid, equally appropriate translations).

Interpretation

Trying to compare BLEU scores across different corpora and languages is strongly discouraged. Even comparing BLEU scores for the same corpus but with different numbers of reference translations can be highly misleading.

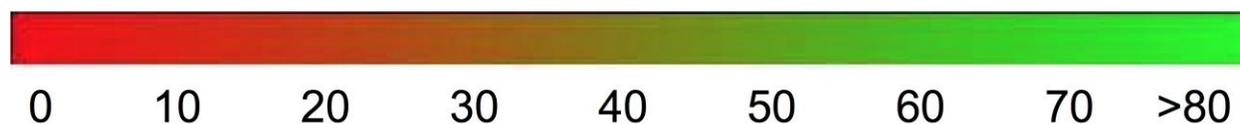
However, as a rough guideline, the following interpretation of BLEU scores (expressed as percentages rather than decimals) might be helpful.

¹ More details on the data sets can be found out e.g. for the 2020 edition on <http://statmt.org/wmt20/translation-task.html>

² Conference website: <https://iwslt.org/>

BLEU Score	Interpretation
< 10	Almost useless
10 - 19	Hard to get the gist
20 - 29	The gist is clear, but has significant grammatical errors
30 - 40	Understandable to good translations
40 - 50	High quality translations
50 - 60	Very high quality, adequate, and fluent translations
> 60	Quality often better than human

The following color gradient can be used as a general scale [interpretation of the BLEU score](#):



Segment Analysis

To determine which translations improved the most between the non-customized and the customized Amazon Translate we ranked the translations by the most improved [COMET](#) metric³.

The table below shows the 30 most improved translations with a visualization of edits that have to be applied to transform the Amazon Translate translation into the Amazon Translate Active Custom Translation.

source	reference	difference
Domestic rabbits	domácí králíci.	KDomácí králíci
Increasing calls at EU level for indicator-based evaluations have been linked to new funding opportunities for comparative indicators and evaluation frameworks.	Stále silnější poptávka po hodnoceních, která vycházejí z ukazatelů, na úrovni EU souvisí s novými možnostmi financování srovnávacích ukazatelů a rámců pro hodnocení.	S novými možnostmi financování srovnávacích ukazatelů a rámců hodnocení na úrovni EU na úrovni EU Zvyšující se výzvy na úrovni EU k hodnocením založeným na ukazatelích souvisěj s novými možnostmi financování srovnávacích ukazatelů a hodnotících rámců rámců hodnocení.
in England and Wales, the High Court of Justice, or in the case of a maintenance judgment, the	v Anglii a Walesu u "High Court of Justice" nebo, jde-li o soudní rozhodnutí ve věcech výživného,	v Anglii a Walesu u „High Court of Justice“ nebo v případě rozsudku ve výživném, jde-li o rozhodnutí ve věcech výživného, u

³ BLEU works only really well on a corpus (entire text) level than on an individual sentence level

source	reference	difference
Magistrate's Court on transmission by the Secretary of State;	u "Magistrate's Court" prostřednictvím "Secretary of State";	„Magistrátate's Court-o předání“ prostřednictvím „Secretary of State“;
For the purposes of implementing this Directive, stocks may be established, under individual agreements between Governments, within the territory of a Member State for the account of undertakings established in another Member State.	Pro účely provádění této směrnice mohou být na základě zvláštních mezivládních dohod vytvářeny zásoby na území jednoho členského státu na účet podniků usazených v jiném členském státě.	Pro účely provádění této směrnice mohou být populace zřízeny na základě jednotlivých dohod mezi vládami na základě zvláštních mezivládních dohod vytvářeny zásoby na území jednoho členského státu na účet podniků usazených v jiném členském státě.
all electricity producers and electricity supply undertakings established within their territory to supply their own premises, subsidiaries and eligible customers through a direct line;	výrobcům elektřiny a dodavatelským podnikům usazeným na jejich území zásobovat své vlastní provozy, dceřiné společnosti a oprávněné zákazníky prostřednictvím přímých vedení;	všichni výrobci elektřiny a podniky dodávající elektřinu teluským podnikům usazeným na jejich území, aby dodávaly své vlastní provozy, dceřiné společnosti a způsobilé oprávněné zákazníky prostřednictvím přímých vedení;
That fraction shall be determined in such a way that the value added tax thus due is, in any event, equal to at least 5 % of the amount determined in accordance with paragraphs 1, 2 and 3.;	Tento podíl se stanoví tak, aby vyměřená daň z přidané hodnoty byla v každém případě rovna nejméně 5 % částky určené podle odstavců 1, 2 a 3.;	Tento zlomek podíl se stanoví tak, aby se takto splatná daň z přidané hodnoty rovnala alespoň 5% částky stanovené v souladu s byla v každém případě rovna nejméně 5% částky určené podle odstavců 1, 2 a 3. ;;
Large purchases would drive up prices in the secondary market, defeating the point of the whole operation.	Mohutné odkupy by zase vyšroubovaly ceny na sekundárním trhu, což by popřelo smysl celé operace.	Velké nákupy by zvýšily ceny na sekundárním trhu a překonaly by bod, což by porazilo smysl celé operace.
Member States may allow or require further retention only after they have carried out a thorough assessment of the necessity and proportionality of such further retention, and where they consider it to be justified as necessary for the prevention, detection or investigation of money laundering or terrorist financing.	Členské státy mohou další uchování osobních údajů povolit nebo požadovat pouze poté, co provedou důkladné posouzení jeho nezbytnosti a přiměřenosti, a pokud shledají, že je odůvodněné, neboť je nezbytné pro předcházení, odhalování nebo vyšetřování praní peněz nebo financování terorismu.	Členské státy mohou další uchování osobních údajů povolit nebo požadovat další uchování pouze až poté, co provedou důkladné posouzení jeho nezbytnosti a přiměřenosti tohoto dalšího zadržetí, a pokud to považují za odůvodněné pro prevenci, a považují-li toto další uchování osobních údajů odůvodněné za nezbytné pro předcházení, odhalování nebo vyšetřování praní špinavých peněz nebo teroristů financování.
contracting authority: the State or the legal person governed by public or private law which concludes the contract, or on behalf of which the contract is concluded, with the contractor,	zadavatel: stát nebo právnická osoba veřejného nebo soukromého práva, která zadává zakázku nebo jejímž jménem se zakázka zhotoviteli zadává,	zadavatelem: stát nebo právnická osoba, která se řídí veřejným nebo soukromým právem, která uzavírá smlouvu dává zakázku nebo jejímž jménem je smlouva uzavřena se zhotovitelem zakázka dodavateli zadává,
Intermediate handling or storage of Category 1 material shall take	Přechodná manipulace nebo přechodné uskladnění materiálů	Zprostředkovatelná manipulace s přechodná manipulace nebo přechodné

source	reference	difference
place only in Category 1 intermediate plants approved in accordance with Article 10.	kategorie 1 smí probíhat pouze v zařízeních pro přechodnou manipulaci kategorie 1 schválených v souladu s článkem 10.	uskladnění materiálů kategorie 1 nebo skladování se provádí pouze v prostředcích zařízeních smí probíhat pouze v zařízeních pro přechodnou manipulaci kategorie 1 schválených v souladu s článkem 10.
Procedure for individual approval	Postup pro jednotlivé schválení	Postup pro individuální schvalování
If the results of the expanded inspection satisfy the Member State in accordance with paragraph 2, the access refusal order must be lifted.	Jestliže výsledky rozšířené inspekce uspokojí členský stát v souladu s odstavcem 2, je příkaz k odepření přístupu zrušen.	Pokud výsledky rozšířené kontroly splňují inspekce vyhovují členskému státu v souladu s odstavcem 2, musí být příkaz k zamítnutí odepření přístupu zrušen.
A positive result shall be positive PCR amplification at the expected size (412 bp), with all negative controls being negative and all positive controls being positive.	Za pozitivní výsledek se považuje pozitivní amplifikace PCR s očekávanou velikostí (412 bp), přičemž všechny negativní kontrolní vzorky musí být negativní a všechny pozitivní kontrolní vzorky pozitivní.	P Za pozitivním výsledkem se považuje pozitivní zesílená amplifikace PCR při očekávané velikosti (412 bp), přičemž všechny negativní kontrolní vzorky musí být negativní a všechny pozitivní kontrolní vzorky pozitivní.
The Community and its Member States are members of the World Trade Organisation and parties to Multilateral Environment Agreements, notably the Convention on Biological Diversity, the Framework Convention on Climate Change and the Convention to Combat Desertification;	Společenství a jeho členské státy jsou členy Světové obchodní organizace a stranami mnohostranných dohod o životním prostředí, zejména Úmluvy o biologické rozmanitosti, Rámcové úmluvy o změně klimatu a Úmluvy o boji proti desertifikaci.	Společenství a jeho členské státy jsou členy Světové obchodní organizace a stranami mnohostranných dohod o životním prostředí, zejména Úmluvy o biologické rozmanitosti, Rámcové úmluvy o změně klimatu a Úmluvy o boji proti rozšiřování pouští a desertifikaci;
shall in no circumstances be used other than for taxation purposes or in connection with judicial proceedings or administrative proceedings involving sanctions undertaken with a view to, or in relation to, the making or reviewing of the tax assessment.	nesmějí být za žádných okolností použity pro jiné účely než pro daňové účely nebo pro účely soudního řízení, trestního řízení nebo řízení o uložení správních sankcí, pokud se týká vyměření daně nebo přezkoumání vyměření daně.	nesmějí být za žádných okolností použity pro jiné účely než pro daňové účely nebo v souvislosti se soudním řízením pro účely soudního nebo správního řízení týkajícím se sankcí za účelem o sankcích, pokud se týká vyměření daně nebo přezkoumání daňového vyhledávání vyměření daně nebo v souvislosti s ním.
The structure of the authorisations for the use of special seals shall follow the structure defined in Title II. in relation with D.E. 1/6 Decision reference number.	Struktura povolení k použití zvláštních závěr se řídí strukturou definovanou v hlavě II ve vztahu k DP 1/6 Referenční číslo rozhodnutí.	Struktura povolení k používání zvláštních plomb závěr se řídí strukturou vymezenou v hlavě II ve vztahu k DP 1/6 Referenční číslo rozhodnutí.
Where persons authorized or entitled to supply medicinal products to the public may also, under national law, engage in	Pokud osoby oprávněné nebo zmocněné vydávat léčivé přípravky veřejnosti mohou podle vnitrostátních právních	Pokud mohou osoby oprávněné nebo oprávněné vydávat léčivé přípravky veřejnosti také mohou rovněž podle vnitrostátních právních předpisů

source	reference	difference
wholesale business, such persons shall be subject to the authorization provided for in paragraph 1.	předpisů také provozovat distribuci, podléhají takové osoby povolení uvedenému v odstavci 1.	<u>provozovat</u> velkoobchodní činnost, podléhají těmto osobám povolení <u>podle stanovenému v odstavci</u> 1.
Only one language can be chosen during deployment creation.	Během vytváření rozmístění lze zvolit pouze jeden jazyk.	Během vytváření <u>nazerozmístění</u> lze zvolit pouze jeden jazyk.
the Kingdom of Belgium, the Kingdom of Denmark, the Federal Republic of Germany, the Hellenic Republic, the Kingdom of Spain, the French Republic, Ireland, the Italian Republic, the Grand Duchy of Luxembourg, the Kingdom of the Netherlands, the Republic of Austria, the Portuguese Republic, the Republic of Finland, the Kingdom of Sweden, the United Kingdom of Great Britain and Northern Ireland.	Belgické království, Dánské království, Spolková republika Německo, Řecká republika, Španělské království, Francouzská republika, Irsko, Italská republika, Lucemburské velkovévodství, Nizozemské království, Rakouská republika, Portugalská republika, Finská republika, Švédské království, Spojené království Velké Británie a Severního Irska.	Belgick ý <u>ým</u> é královstvím, Dánsk ý <u>ým</u> é královstvím, Spolkov ou <u>ou</u> á republikou, Německo, Řeck ou <u>ou</u> á republikou, Španělsk ý <u>ým</u> é královstvím, Francouzsk ou <u>ou</u> á republikou, Irsk o <u>o</u> , Italsk ou <u>ou</u> á republikou, Lucembursk ým <u>ým</u> é velkovévodstvím, Nizozemsk ým <u>ým</u> é královstvím, Rakousk ou <u>ou</u> é republikou, Portugalsk ou <u>ou</u> é republikou, Finsk ou <u>ou</u> é republikou- <u>y</u> Švédské království, Spojené království Velké Británie a Severního Irska.
Where the undertaking is subject to collective insolvency proceedings or fulfils the criteria under its domestic law for being placed in collective insolvency proceedings at the request of its creditors.	je-li podnik předmětem kolektivního úpadkového řízení nebo splňuje kritéria stanovená ve vnitrostátních právních předpisech pro to, aby vůči němu bylo na žádost jeho věřitelů zahájeno kolektivní úpadkové řízení;	<u>Pokud je podnik předmětem kolektivního insolvenčního úpadkového řízení nebo splňuje kritéria podle svého stanovená ve vnitrostátního práva pro zahájení kolektivního insolvenčního řízení předpisech pro to, aby vůči němu bylo na žádost svých jeho věřitelů zahájeno kolektivní úpadkové řízení.</u>
Three-pillar system	Systém tří pilířů	Třípilíř ový <u>ový</u> systém
Under reserve of the entry into force of the Treaty concerning the accession of the Republic of Finland to the European Union.	S výhradou vstupu v platnost smlouvy o přistoupení Finské republiky k Evropské unii.	V <u>rámci rezervy</u> S výhradou vstupu v platnost <u>S</u> smlouvy o přistoupení Finské republiky k Evropské unii.
Supporters of the Euro point to the success of the US, with its single currency.	Stoupenci eura poukazují na úspěch USA, které mají jedinou měnu.	<u>Přízniv</u> Stoupenci eura <u>po</u> ukazují na úspěch USA, s jednotnou měnou.
the measures decided concerning monitoring in accordance with Article 10.	přijátá opatření týkající se monitorování v souladu s článkem 10.	opatření, <u>o nichž se rozhodlo přijatá</u> v souvislosti s <u>kontrolováním</u> v souladu s článkem 10.
This is a very time consuming operation but, when used in conjunction with high antialiasing and soft shadows it can produce very high quality renderings with fewer radiosity artifacts.	Toto je velmi časově náročná operace, ale pokud se provede ve spojení s vysokou úrovní antialiasingu a jemnými stíny, může vytvářet rendrování o velmi vysoké kvalitě s méně artefakty radiosity.	<u>Jedná se o</u> Toto je velmi časově náročn ou <u>ou</u> á operac ie <u>ie</u> , ale <u>při použití</u> <u>okud se provede</u> ve spojení s vysok ým <u>ým</u> vyhlazováním a <u>měk</u> čkou úrovní antialiasingu a <u>jemnými</u> stíny, může <u>produkovat velmi kvalitní</u> rendrování s <u>menším počtem</u> vytvářet <u>rendrování o velmi vysoké kvalitě s</u> <u>několika</u> artefakt ů <u>ů</u> radiosity.

source	reference	difference
Chapter 7: New and emerging drug trends and action on new drugs	Kapitola 7: Trendy užívání nových a nově se rozmáhajících drog a opatření proti novým drogám	Kapitola 7: Nové a vznikající Trendy užívání nových a nově se rozmáhajících drog ové trendy a akce na a opatření proti novým drogám
The annual payments of the State and the City are covered by their respective budgets.	Roční platby státu a města jsou hrazeny z jejich příslušných rozpočtů.	Roční platby státu a města jsou kryty jejich hrazeny z jejich příslušných rozpočtů.
Where a currency conversion service is offered prior to the initiation of the payment transaction and where that currency conversion service is offered at an ATM, at the point of sale or by the payee, the party offering the currency conversion service to the payer shall disclose to the payer all charges as well as the exchange rate to be used for converting the payment transaction.	Je-li služba konverze měny nabídnuta před iniciováním platební transakce a je-li tato služba konverze měny nabídnuta v bankomatu, místě prodeje nebo příjemcem, strana, jež plátcí službu konverze měny nabízí, sdělí plátcí veškeré poplatky, jakož i směnný kurz, který bude při konverzi u platební transakce použit.	Pokud je služba převodu konverze měny nabízena před zahájením platební transakce a je-li pokud je tato služba převodu konverze měny nabízena v bankomatu, v místě prodeje nebo příjemcem, strana, kte rá plátcí nabízí službu převodu měny plátcí , sdělí plátcí také všechny rovněž veškeré poplatky, jako směnný kurz, který semá být použit pro převod platební transakce.
Opinion of 18 April 2006 on a notification for prior checking concerning the 'Attestation procedure' dossier (Case 2006-92)	Stanovisko ze dne 18. dubna 2006 k oznámení v souvislosti s předběžnou kontrolou spisu „Osvědčovací řízení“ (věc 2006-92)	Stanovisko ze dne 18. dubna 2006 k oznámení ev souvislosti s předběžnou kontrolou le týkající se dokumentace „Potvrzou spisu „Schválení““ (věc 2006-92)
In applying the provisions of paragraph 1, the Commission may, inter alia:	Při používání ustanovení odstavce 1 komise může kromě jiného:	Při uplatňování odstavce 1 může Komise mimo jiné:

Results Summary

By customizing Amazon Translate Active Custom Translation with TAUS Financial domain English-Czech training data the test set BLEU score improves by almost 6 points. This is a significant improvement in machine translation quality for this language pair and domain.

This numerical score improvement is supported by analyzing segments that improved significantly according to the semantic quality measure COMET: translations became more fluent and adhered better to the terminology used in the Financial domain.

Polyglot Technology

Polyglot Technology LLC helps customers succeed with machine translation by ensuring that they make best use of data available to them, by assessing machine translation quality independent from MT vendors and by advising customers on how to best integrate the technology with people and processes.