

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

Language Pair

English → Romanian

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

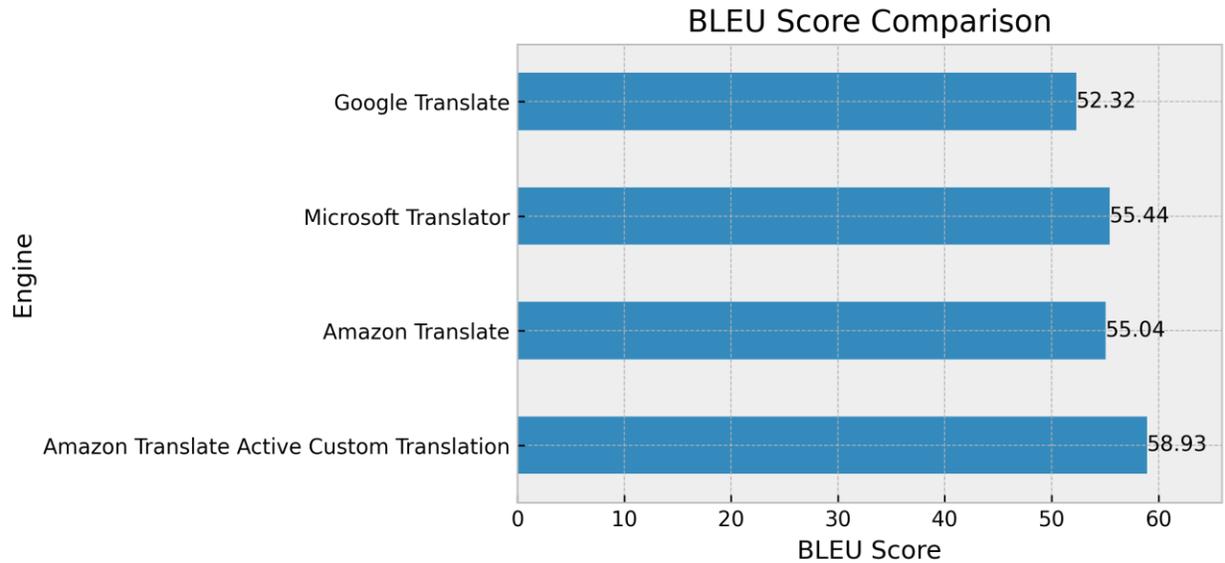
To judge whether machine translation is good or not, human evaluation is the best method. We can ask speakers of the source language and the target language, or better professional translators, to judge whether a machine translation is an adequate and fluent translation of the original text. Or we can 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.

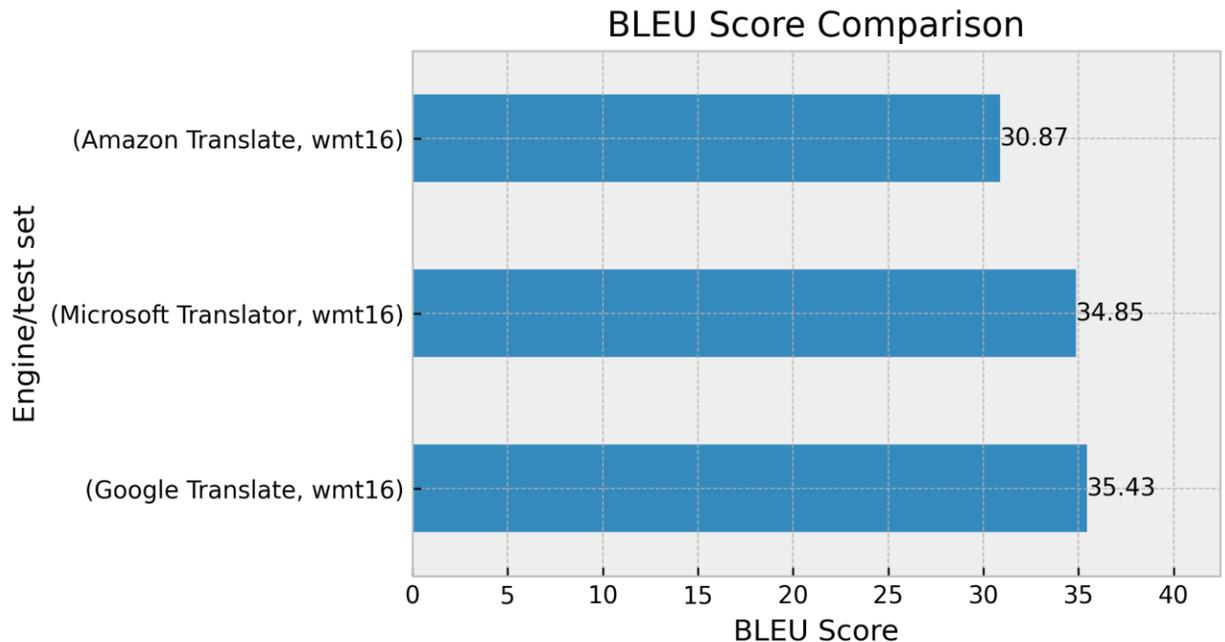
TAUS Test Set

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.

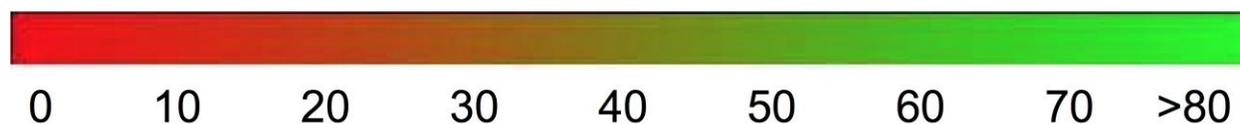
However, as a rough guideline, the following interpretation of BLEU scores (expressed as percentages rather than decimals) might be helpful.

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
For the components Loss absorbing capacity ('LAC') of technical provisions and/or deferred taxes when reported as a separate component it shall be the amount of the loss-absorbing capacity (these amounts shall be reported as negative values)	Pentru componentele Capacitatea rezervelor tehnice și/sau a impozitelor amânate de a absorbi pierderile raportate ca o componentă separată, se raportează cuantumul aferent capacității de absorbție a pierderilor (aceste cuantumuluri trebuie raportate ca valori negative)	Pentru componentele Capacitatea de absorbție a pierderilor („LAC”) a provizion rezervelor tehnice și/sau a impozitelor amânate atunci când sunt de a absorbi pierderile raportate ca o componentă separată, valoare ase raportează cuantumul aferent capacității de absorbție a pierderilor (aceste valoc uantumuluri se raportează ca valori negative)
By contrast, since the School Transport Scheme is an existing aid	În schimb, întrucât schema de transport școlar este o schemă de	În schimb, întrucât schema de transport în școli este o schemă de ajutor

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

source	reference	difference
scheme which Ireland continues to operate, it is in the context of the dialogue referred to in the previous recital that the Commission intends to inform Ireland why it believes that that scheme is incompatible with the internal market.	ajutor existentă pe care Irlanda continuă să o utilizeze, în contextul dialogului menționat la considerentul anterior Comisia intenționează să informeze Irlanda cu privire la motivele pentru care consideră că schema respectivă este incompatibilă cu piața internă.	existentă pe care Irlanda continuă să o funcționeze, <u>Comisia intenționează să informeze Irlanda</u> în contextul dialogului menționat în considerentul anterior, <u>Comisia intenționează să informeze Irlanda de ce consideră că schema respectivă este incompatibilă cu piața internă.</u>
This is a major problem because qualifications are granted by the education systems of EU Member States, of which there are at least 27.	Aceasta este o problemă majoră, deoarece calificările sunt acordate de sistemele educaționale din statele membre UE, care sunt cel puțin 27 la număr.	Aceasta este o problemă majoră, deoarece calificările sunt acordate de sistemele educaționale ale statelor membre ale UE, dintre care există cel puțin 27 <u>de ani.</u>
Small pelagic fisheries in the Aegean Sea and Crete Island	Activități de pescuit de specii pelagice mici în Marea Egee și în insula Creta	Pescuire <u>teul de specii</u> pelagice mici în Marea Egee și <u>în</u> insula Creta
The number of Claims open at the beginning of the year and closed at the end of the year and settled with payments, by accident/underwriting years from the year N-1 (the year before the reporting year) to N-14, amount of all previous reporting periods prior to N-14 and the total of all the years from N-1 to prior to year N-14.	Numărul de daune deschise la începutul anului care sunt închise la sfârșitul anului și soluționate cu plăți, pe ani de accident/subscriere, de la anul N-1 (anul care precedă anul de raportare) până la anul N-14, cuantumul aferent tuturor perioadelor dinainte de anul N-14 și totalul pentru toți anii de la N-1 până la perioadele dinainte de N-14.	<u>Numărul de Revendicări daune</u> deschise la începutul anului <u>șicare sunt</u> închise la sfârșitul anului și <u>desoluționate</u> cu plăți, pe <u>ani de</u> accident/ <u>ani de</u> subscriere, <u>dine la</u> anul N-1 (anul <u>dinaintea care precedă</u> anul de raportare) până la <u>N-14, valoarea tuturor perioadelor anterioare de raportare anterioare anului N-14, cuantumul aferent tuturor perioadelor dinainte de anul N-14 și totalul tuturor anilor pentru toțianii</u> de la N-1 <u>până la perioada anteriori înară la</u> anul N-14.
Box reference I.13: In the case animals for slaughter either Assembly centre or Establishment shall be marked as a Place of destination as described in Article 7 of Directive 64/432/EEC.	Rubrica de referință I.13: În cazul animalelor pentru sacrificare, se marchează ca Loc de destinație fie Centrul de colectare, fie Unitatea, astfel cum se prevede la articolul 7 din Directiva 64/432/CEE.	Rubrica de referință I.13: În cazul animalelor <u>destinate sacrificării, pentru sacrificare, se marchează ca Loc de destinație</u> fie <u>Centrul de adunare</u> colectare, fie <u>Unitatea</u> se marchează ca <u>loc de destinație conform, astfel cum se prevede la articolul</u> 7 din Directiva 64/432/CEE.
These are the technical provisions for non-proportional health reinsurance, without risk margin after deduction of the amounts recoverable from reinsurance contracts and SPVs, with a floor equal to zero, relating to life activities.	Se raportează rezervele tehnice pentru reasigurarea neproportională de sănătate, fără marja de risc după deducerea sumelor recuperabile din contracte de reasigurare și de la vehicule investiționale, cu un prag egal cu zero, referitoare la activitățile de asigurare de viață.	<u>Acestea sunt provizioan</u> Se raportează <u>rezervele</u> tehnice pentru reasigurarea <u>sanitară</u> neproportională <u>de sănătate</u> , fără <u>marjă</u> de risc după deducerea sumelor recuperabile din contractele de reasigurare și <u>SPV de la vehicule investiționale</u> , cu un <u>nivel prag</u> egal cu zero, <u>a referenteitoare la</u> activitățile <u>de asigurare</u> de viață.

source	reference	difference
It is not Parliament that you should be putting pressure on, but the European Council.	Nu asupra Parlamentului ar trebui să faceți presiuni, ci asupra Consiliului European.	Nu <u>asupra</u> Parlamentului <u>ar</u> trebui să faceți presiuni <u>asupra acestuia</u> , ci asupra Consiliului European.
in the case of a Cash Value Insurance Contract or an Annuity Contract, by the Financial Institution that is obligated to make payments with respect to the contract.	în cazul unui Contract de asigurare cu valoare de răscumpărare sau al unui Contract cu rentă viageră, de Instituția financiară care este obligată să facă plăți cu privire la respectivul contract.	în cazul unui <u>eContract</u> de asigurare <u>pe</u> valoare <u>în nume de răscumpărare</u> sau al unui <u>eContract de anuitate, de către</u> <u>icu rentă viageră, de</u> Instituția <u>f</u> financiară care este obligată să <u>efectueze</u> <u>facă</u> plăți în legătură cu contractul <u>respectiv</u> .
It does not avoid taxation.	Aceasta nu evită impozitarea.	<u>N</u> Aceasta <u>nu</u> evită impozitarea.
International contracted staff shall have the nationality of a Member State.	Membrii personalului internațional contractual trebuie să dețină naționalitatea unui stat membru.	<u>Personalul internațional contractat are</u> <u>cețățeni</u> <u>Membrii personalului contractat la nivel internațional trebuie să dețină naționalitatea</u> a unui stat membru.
Fully Automated Storage Tiering (FAST): Designed to enable more efficient storage resource utilization by analyzing the activity of data with a LUN.	Organizare complet automată pe niveluri pentru stocare (FAST): creată pentru a permite utilizarea mai eficientă a resurselor de stocare prin analizarea activității datelor cu o unitate LUN.	<u>CFAST (Creare complet A</u> <u>automated Storage Tier (FAST): Proiectă de niveluri pentru stocare): creat</u> pentru a permite <u>e</u> <u>utilizarea</u> mai eficientă a resurselor de stocare prin analizarea activității <u>ttii</u> datelor cu un LUN.
issues agreed upon by the Parties, such as decisions adopted, statements agreed upon and any conclusions.	chestiunile asupra cărora părțile au ajuns la un acord, precum deciziile adoptate, declarațiile convenite și eventualele concluzii.	<u>aspecte</u> <u>conve</u> <u>chestiuni</u> <u>le de</u> <u>părți</u> <u>asupra cărora părțile au ajuns la un</u> <u>acord</u> , cum ar fi deciziile adoptate, declarațiile convenite și orice concluzii.
Total amount of other contributions, outside ESI Funds paid to the financial instrument (in EUR)	Suma totală a altor contribuții, în afara fondurilor ESI, făcute în beneficiul instrumentului financiar (în EUR)	<u>Valoare</u> <u>Suma</u> totală a altor contribuții, în afara fondurilor ESI, <u>plătite în</u> <u>beneficiul</u> instrumentului financiar (în EUR)
This is the absolute value of the assets sensitive to health disability — morbidity risk — Medical expense charge due to an increase of medical payments, before the shock.	Se raportează valoarea absolută a activelor senzitive la cerința pentru riscul de invaliditate-morbiditate în asigurarea de sănătate – cheltuieli medicale, ca urmare a unei creșteri a plăților medicale, înainte de șoc.	<u>Aceasta este</u> <u>Se raportează</u> valoarea absolută a activelor <u>sensibile la</u> <u>dezabilitatea sănătății —</u> <u>zitive la cerința pentru riscul de morbiditate —</u> <u>Taxa de invaliditate-morbiditate în asigurarea de sănătate —</u> cheltuieli medicale <u>datorată, ca urmare a</u> unei creșteri a plăților medicale, înainte de șoc.
The measure consists of establishing a Charter Fund Scheme (the 'Charter Fund' or the 'Scheme') which will grant aid to tour operators flying air charter flights [8] to three counties in Northern Norway:	Măsura constă în instituirea unei scheme de fond pentru zboruri charter („fondul pentru zboruri charter” sau „schema”) care va acorda ajutor agențiilor de turism care gestionează zboruri charter [8] către trei districte din nordul Norvegiei:	Măsura constă în instituirea unei scheme <u>a</u> <u>F</u> <u>de</u> <u>fondu</u> <u>luri</u> pentru <u>E</u> <u>chartăer</u> („ <u>F</u> <u>fondul pentru E</u> <u>chartăer</u> ” sau „schema”) care va acorda ajutor operatorilor de turism care efectuează zboruri charter aeriene [8] către trei județe din nordul Norvegiei:

source	reference	difference
Financial corrections by the Joint Managing Authority	Corecțiile financiare efectuate de către autoritatea de gestionare comună	Corecții financiare de către efectuate de autoritatea comună de management gestionare
Cambodia has accepted all the recommendations of the United Nations Human Rights Council in the Universal Periodic Review.	Cambodgia a acceptat toate recomandările Consiliului Organizației Națiunilor Unite pentru Drepturile Omului în cadrul evaluării periodice universale.	Cambodgia a acceptat toate recomandările Consiliului ONU pentru Drepturile Omului în Rcadrul revizuirea Pii periodică de universalăe.
To obtain all the authorizations, permits and agreements necessary for conducting the works related to the connection installation, obtain the agreements necessary for regulating the use rights of the land for the energy capacities that will enter the Operator’s domain, the accord of land owners;	Obținerea tuturor autorizațiilor, avizelor și acordurilor necesare executiei instalației de racordare, a contractelor necesare pentru reglementarea dreptului de folosință a terenului pentru capacitățile energetice ce vor intra în patrimoniul Operatorului, a acordurilor proprietarilor de teren;	Pentru o Obținerea tuturor autorizațiilor, autorizații vizele și acordurilor necesare pentru desfășurarea lucrărilor afere ntele executiei instalației de racordare, obținerea acordurilor contractelor necesare pentru reglementăriiarea drepturilor de utilizare ului de folosință a terenului fi orului pentru capacitățile energetice care vor intra în domeniul patrimoniului Operatorului, a acordurilor proprietarilor de terenuri;
The normal value is the actual domestic price per product type of only the profitable domestic sales of the product types during the AIP, if:	Valoarea normală este prețul intern real pe tip de produs numai pentru vânzările interne rentabile ale tipurilor de produse în cursul PAAM, în cazul în care:	Valoarea normală este prețul intern real pe tip de produs numai pentru vânzările interne prof rentabile ale tipurilor de produse în tim p cursul API, dacă în cazul în care:
Capital requirement for other financial sectors (Non–insurance capital requirements) — Institutions for occupational retirement provisions	Cerințe de capital pentru alte sectoare financiare (cerințe de capital fără legătură cu asigurările) — instituții pentru furnizarea de pensii ocupaționale	Cerințae de capital pentru alte sectoare financiare (cerințe de capital, alte decât cele de fără legătură cu asigurările) — in stituții pentru proviz oane furnizarea de pensii ocupaționale
AC, Amending Corrigendum concerning Standard Documents other than IMO.	AC, Rectificare cu privire la alte Documente Standard decât cele ale OMI.	AC, Rectificare de modificare arectificativă privind alte documentelor standard, alte decât OMI.
Amount/Estimate of the overall adjustment for loss–absorbing capacity for deferred taxes, including the part embedded in each component and the part reported as a single component.	Valoarea/estimarea ajustării globale pentru capacitatea impozitelor amânate de a absorbi pierderile, incluzând partea inclusă în fiecare componentă și partea raportată ca o componentă separată.	Suma sau Valoarea/estimarea ajustării globale pentru capacitatea de impozitelor amânate de a absorbi ie pierderilor pentru impozitele amânate, inclusivzând partea și in corporată usă în fiecare componentă și partea raportată ca o singură componentă separată .
The structure of a third country unique identification number shall follow the structure defined in Title II for D.E.	Structura numărului de identificare unic acordat de o țară terță trebuie să respecte structura definită în titlul II pentru E.D.	Structura unui număr unui numărului de identificare al unei țări unic acordat de o țară terță ea trebuie să respecte structura definită în titlul II pentru D-E-E D.
The accounting for the acquisition of an interest in such a joint	Contabilizarea dobândirii unui interes într-o astfel de exploatare	Contabilitatea dobânzii pedirii unei participații într-o astfel de operațiune

source	reference	difference
operation is specified in paragraphs B33A–B33D.	în participație este prevăzută la punctele B33A-B33D.	în comun este specificată la punctele B33A-B33D.
The indicative compendium of environmental goods and services and of economic activities is established as set out in the Annex.	Culegerea cu titlu indicativ a bunurilor și serviciilor de mediu și a activităților economice este instituită așa cum se prevede în anexă.	Compendiululegerea orientativă al bunurilor și serviciilor de mediu și al activităților economice seeste stabilește în conformitate cuită astfel cum se prevede în anexă.
The MCR is subject to an absolute floor depending on the nature of the undertaking (as defined in Article 129 (1) (d) of the Directive 2009/138/EC).	La stabilirea MCR se aplică un prag absolut în funcție de natura societății [astfel cum este definit la articolul 129 alineatul (1) litera (d) din Directiva 2009/138/CE].	MCR face obiectul unui nivelLa stabilirea MCR se aplică un prag absolut în funcție de natura întreprinderii [astfel cum este definită la articolul 129 alineatul (1) litera (d) din Directiva 2009/138/CE].
Such calculation is likely to involve an unjustified administrative burden in relation to the effect of these transactions on Spain's total VAT own resources base.	Este probabil ca astfel de calcule să implice o sarcină administrativă nejustificată pentru Spania în raport cu efectul acestor operațiuni asupra bazei totale a resurselor proprii din TVA ale acesteia.	UnEste probabil ca astfel de calcul peate să implice o sarcină administrativă nejustificată pentru Spania în raport cu efectul acestor operațiuni asupra bazei totale a resurselor proprii din TVA a Spaniile acesteia.
Should companies from overseas be investing huge amounts of money into sporting clubs within Europe?	Ar trebui să permitem companiilor de dincolo de ocean să investească sume enorme de bani în cluburile sportive din Europa?	În cazul înAr trebui care companiile din străinătate să investească sume uriașe de bani în cluburile sportive din Europa?
The price comparison was made on a type-by-type basis for transactions at the same level of trade, duly adjusted where necessary, and after deduction of rebates and discounts.	Compararea prețurilor a fost realizată per tip de produs pentru tranzații la același nivel de comercializare, ajustate în mod adecvat acolo unde a fost necesar, și după deducerea rabaturilor și a reducerilor.	Comparația prețurilor a fost efectuată pe bază de tip de produs pentru tranzațiile efectuate la același nivel de comercializare, ajustate corespunzător, după caz, acolo unde a fost necesar și după deducerea rabaturilor și a reducerilor.

Results Summary

By customizing Amazon Translate Active Custom Translation with TAUS Financial domain English-Romanian training data the test set BLEU score improves by close to 4 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.