

Bayona, Pamela

Working Paper

Beyond six digits: Automated tariff line HS transposition using Natural Language Processing

WTO Staff Working Paper, No. ERSD-2025-04

Provided in Cooperation with:

World Trade Organization (WTO), Economic Research and Statistics Division, Geneva

Suggested Citation: Bayona, Pamela (2025) : Beyond six digits: Automated tariff line HS transposition using Natural Language Processing, WTO Staff Working Paper, No. ERSD-2025-04, World Trade Organization (WTO), Geneva

This Version is available at:

<https://hdl.handle.net/10419/314422>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

World Trade Organization

Trade Policies Review Division

**BEYOND SIX DIGITS:
AUTOMATED TARIFF LINE HS TRANSPOSITION
USING NATURAL LANGUAGE PROCESSING**

Pamela Bayona
*World Trade Organization*¹

Manuscript date: 5 November 2024

Disclaimer: This is a working paper, and hence it represents research in progress. This paper represents the opinions of individual staff members or visiting scholars and is the product of professional research. It is not meant to represent the position or opinions of the WTO or its Members, nor the official position of any staff members. Any errors are the fault of the author.

¹ Pamela Bayona is Statistical Officer at the Trade Policies Review Division of the World Trade Organization. Email: pamela.bayona@wto.org.

Beyond Six Digits: Automated Tariff Line HS Transposition Using Natural Language Processing

Pamela Bayona²

Trade Policies Review Division
World Trade Organization

Manuscript Date: 5 November 2024

ABSTRACT

This paper explores the application of Natural Language Processing (NLP) techniques to automate Harmonized System (HS) tariff line transposition, employing a three-stage process: unique 1:1 tariff code matching (Round 1), exact description matching (Round 2), and “smart” description matching (Round 3) using Artificial Intelligence (AI) and lexical similarity methods paired with harmonized 6-digit concordance and cosine similarity. Similarity is calculated using either Term Frequency Inverse Document Frequency (TF-IDF) vectors or Sentence-BERT (SBERT) embeddings, comparing two scenarios: a straightforward case (Economy A) with standardized descriptions, and a complex case (Economy B), with more detailed technical descriptions. Results indicate that automated HS transposition can significantly augment the efficiency of traditionally manual methods, reducing processing time from two to three weeks to approximately half a day (up to 30 times faster). The overall accuracy rate is 99.6% for the simpler scenario and 98.8% for the complex one, for a standard set of approximately 10,000 HS codes. While non-AI techniques cover most of the accurate matches, AI-based Round 3 techniques address cases requiring the most manual effort. SBERT generally outperforms TF-IDF, however including subheadings tends to reduce its accuracy. In certain cases, particularly for highly technical tariffs, TF-IDF's straightforward approach provides an advantage over SBERT. Overall, NLP techniques hold significant potential for improving HS transposition methods and facilitating the development of richer tariffs and trade datasets to enable more in-depth analyses. Future research should focus on refining these techniques across diverse datasets to optimize their broader application in tariff and trade data analysis.

Keywords: Harmonized System, tariff line, HS transposition, correlation tables, concordance, natural language processing

JEL Classifications: F10, F13

Disclaimer: This is a working paper, and hence it represents research in progress. This paper represents the opinions of individual staff members or visiting scholars and is the product of professional research. It is not meant to represent the position or opinions of the WTO or its Members, nor the official position of any staff members. Any errors are the fault of the author.

² Gratitude is extended to Rowena Cabos for lending her expertise and extraordinary patience throughout this initiative; Alexander Keck and Maria Donner-Abreu for their guidance and support; Rohini Acharya for earlier endorsing the inception of this pilot; as well as colleagues Alya Belkhodja, Phillipe Chauvet, Arti Gobind Daswani, Antonia Diakantoni, Verena Hess-Bays, Jung-Ah Kang, Takako Ikezuki, Mun How Mong, Jeevanthi Liyana Pathirana, Roy Santana, Eric Ng Shing, Marcelo Vega Soriano, and Dayong Yu for their helpful comments. Finally, I also wish to thank Monica Lezama, Vanessa Bocquet and the Learning Service team for always supporting our staff training requests, which directly contribute to our capacity to undertake this type of work.

EXECUTIVE SUMMARY

The comparison of tariffs across time poses significant challenges when data are expressed in different versions of the Harmonized System (HS). From a practical point of view, such comparison requires that one of the datasets in one version of the HS be "transposed" into the version of the second dataset. However, transposing tariff lines beyond the 6-digit level in trade statistics is complex and time-consuming, often requiring experts to manually match product descriptions across different HS nomenclatures. In the context of the work of the WTO's Trade Policies Review Division (TPRD), and in particular the preparation of Factual Presentations and Trade Policies Reviews under the purview of the Transparency Mechanism, this process is essential for comparing most-favoured-nation (MFN), preferential, and bound tariffs. It can also be relevant during trade negotiations involving older tariff lists. With national schedules containing thousands of detailed tariff line codes, variations across WTO Members and HS versions add to the challenge. Many researchers avoid this detailed transposition due to its technical demands and potential for errors, instead opting for analysis at the less detailed harmonized 6-digit level for convenience.

This paper argues that new artificial intelligence (AI) technologies can automate and streamline HS transposition to a high degree of accuracy. Natural Language Processing (NLP) techniques are applied to **automate tariff line Harmonized System (HS) transposition** using three staged concordance rounds in a single end-to-end process: a unique 1:1 tariff code matching (Round 1), an exact description matching (Round 2), and a "smart" description matching based on harmonized 6-digit level concordance combined with cosine similarity rankings generated from either Term Frequency Inverse Document Frequency (TF-IDF) vectors or Sentence-BERT (SBERT) embeddings (Round 3).

To account for the wide variation in tariff complexity across economies, two tariff scenarios were considered: a "simple" tariff scenario with standardized product descriptions (Economy A)³, and a "complex" tariff scenario with highly technical descriptions (Economy B). To evaluate the usefulness of subheadings, two SBERT versions were tested: one with subheadings and one without. All results were then evaluated by a manual human line-by-line verification.

Key findings include the following⁴:

1. Automated HS transposition can reduce tariff processing time from two to three weeks to half a day (up to 30 times faster), with accuracy rates of 99.6% for simpler tariff scenarios and 98.8% for complex ones, out of a standard set of approximately 10,000 tariff codes.
2. Non-AI automation techniques (Rounds 1 and 2) already accurately cover the majority of tariff matches, achieving 97.8% and 93.8% accuracy for Economies A and B, respectively.
3. AI-based methods (Round 3) cover a lower percentage (2.2% and 6.2% of the codes in Economy A and Economy B, respectively), however they are especially valuable in handling the remaining complex cases that require the most manual work for statisticians.
4. At the minimum, approximately 80% of all Round 3 "smart" matches are correct.
5. Final verification after Round 3 is needed only for codes that do not match their nearest 6-digit MFN average tariff; these account for 1-5% of all tariff codes.
6. SBERT generally outperforms TF-IDF, however the inclusion of subheadings does not provide a net benefit to SBERT performance. Further testing is recommended.
7. The particular syntax and specificity of tariff product descriptions impacts the performance of TF-IDF and SBERT, and in certain cases of highly technical tariffs TF-IDF showed advantages due to its straightforward matching algorithm.

The end result is a fully functional end-to-end automated HS transposition process requiring minimal human verification as compared to traditional processes.

³ Economy names are anonymized.

⁴ This study examined the HS 2017 to HS 2022 transposition for Economy A and Economy B, noting that results may vary depending on the specific complexity of each economy's tariff regime.

The study highlights that incorporating NLP techniques into HS transposition processes offers substantial potential to enhance the efficiency and accuracy of tariff analysis, making it an invaluable tool for trade statisticians. However, it is essential to maintain human oversight to ensure quality and reliability within AI-assisted HS transposition workflows, as highlighted in Round 3 results.

This paper details a novel approach to automating a traditionally manual, labour-intensive and error-prone process in international trade statistics. It highlights some key features comparing the suitability of TF-IDF vectorization and SBERT embeddings methods to enable "smart" tariff description matching. It also discusses results in the specific context of the WCO HS nomenclature and standard WTO methodologies for tariff processing.

For future work, further refinement and testing on diverse datasets is recommended to optimize these methods for broader application. In terms of application, tariff line analysis done to compare MFN duties with either bound duties, preferential duties or MFN duties in another nomenclature could potentially benefit. Additionally, tariffs and trade analysis typically conducted at the 6-digit level could be expanded to the more granular tariff line level, enabling the creation of richer datasets that allow for deeper analyses.

Contents

I. INTRODUCTION.....	1
II. MOTIVATION.....	1
III. LITERATURE REVIEW.....	2
IV. METHODOLOGY	4
A. Overview of the process.....	4
B. Pre-processing of tariff files	5
C. HS transposition	7
D. Note on subheadings.....	9
E. Results evaluation	9
V. RESULTS.....	10
1. Overview of main results.....	10
2. Detailed results for a simple scenario.....	11
3. Detailed results for a complex scenario	14
3.1 Both correct or valid matches.....	15
3.2 One or both incorrect matches	16
4. Computational efficiency	18
VI. TARIFF VERIFICATION	19
1. HS 6-digit MFN average tariff as a benchmark for accuracy	19
2. Wisdom of the crowds as a benchmark for accuracy	19
3. Cosine similarity scores as a benchmark for accuracy	20
VII. CONCLUSIONS	20
VIII. RECOMMENDATIONS	22
IX. BIBLIOGRAPHY	23

I. INTRODUCTION

Tariff line level Harmonized System (HS) transposition⁵ beyond the harmonized 6-digit level⁶ has long been a labor-intensive process in trade statistics. It enables granular comparisons between most-favoured-nation (MFN), preferential, and bound tariffs that exist in different HS nomenclatures. A common example occurs during trade agreement negotiations, which often take place over extended periods of time. As a result, the originally negotiated list of preferential tariffs may only be available in an earlier HS nomenclature, making it difficult to directly compare with the most recent MFN tariffs without the use of a national correlation table provided by the relevant economy.

Traditional manual HS transposition typically involves detailed line by line product description matching done by tariff experts across different versions of the HS tariff nomenclature. This process can take days if not weeks, depending on the volume and complexity of the economy's national tariff lines and the availability of full or partial national correlation tables. There is also a wide variety in the complexity of tariff line level product descriptions across economies and HS nomenclature versions. The latest HS nomenclature (HS 2022) contained 5,612 tariff codes at the 6-digit subheading level; most tariff line level schedules contain between 6,000 to 12,000 tariff codes. Certain 6-digit subheadings, particularly in the chemicals and automotive HS chapters, contain up to nearly 40 different tariff line codes each, though most average around eight codes. Some economies may also produce different tariff schedules within a single year, such that their MFN and preferential tariffs issued within the same year may have different counts and classifications. Others may have much less variations between tariff lines across different nomenclatures, such that their count and classification remain identical or contain only minor changes. Finally, some economies may produce tariffs extending beyond the standard 8-digit, and instead contain 9-, 10-, 11-, 12-digit (or more) levels of detail, as well as multiple tariff line suffixes (TLS) or child lines beyond the original parent tariff line (TL).

For these reasons, most researchers do not attempt tariff line HS transposition and instead choose to maintain analysis at the harmonized 6-digit level, in effect losing granularity in favour of convenience. Aside from being time-consuming, HS transposition can also be error-prone, requiring deep technical chapter-specific knowledge of the history and evolution of the HS nomenclature.

II. MOTIVATION

The regular work of the Regional Trade Agreements (RTA) section in the Trade Policies Review Division (TPRD) of the World Trade Organization is one among several statistical work streams that regularly require tariff line HS transposition. The section uniquely compares preferential tariffs across the years covered by a trade agreement schedule with MFN tariffs in the year of entry into force of the agreement. In the cases when economies' submitted tariffs data overlap different HS nomenclatures, they need to be aligned in a common HS nomenclature before any further data analysis can be performed.

⁵ **HS transposition** refers to the process of updating tariff schedules, trade agreements, and regulatory documents to align with a new version of the Harmonized System ([WTO, 2023](#)). World Customs Organization (WCO) **correlation tables** refer to the tables produced to perform this process since the original 1992 to 1996 HS transposition ([WTO, 1995](#)). "**Concordance**" is a more general term used by statisticians or academics, often referring to tables that align classifications across different classification systems. This terminology comes from the need to *correlate* or *concord* (i.e., bring into agreement) codes from one version or system to another. In actual practice among statisticians, the term "HS transposition" is often used interchangeably with "HS concordance" or "HS correlation".

⁶ The term "**tariff**" (also used interchangeably with [customs or import] "*duty*", "*rate*" or "*duty rate*" in common parlance) is defined as a tax on imports or exports ordinarily levied by governments when a good is imported or exported between economies. Tariffs data is harmonized globally under the Harmonized Commodity Description and Coding System, generally referred to as the "**Harmonized System**" or "**HS**". This multipurpose international product nomenclature is used by more than 200 economies and comprises more than 5,000 commodity groups, identified by a standard 6-digit level code across 99 Chapters. A tariff code, or more commonly "**HS code**", is also referred to as a "*tariff line*". The WCO revises the HS nomenclature to a new version every five years, starting from HS 1992, HS 1997, and so forth until the most recent versions of HS 2012, HS 2017, and HS 2022. Consequently, an HS 6-digit level correlation table is released at the same time between the previous and new version ([WCO, 2024](#)). Standard classifications refer to a "**chapter**" as the 2-digit level, a "**heading**" as the 4-digit level, and a "**subheading**" as the 6-digit level ([WCO, 2018](#)). On the other hand, **national ("tariff-line") level HS codes** are maintained by each economy at the 8, 9, 10, or more levels and can only be officially linked back to the harmonized 6-digit level through the use of national HS correlation tables.

How is tariff line HS transposition used in the regular work of the WTO?

For example, the requisite [WTO Factual Presentation](#) was considered at the WTO Committee for Regional Trade Agreements (CRTA) for Peru's Accession to the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP). For this Agreement, the year of entry into force was 2021 (covered by HS 2017). However many of the CPTPP parties had earlier submitted their preferential tariffs in 2018, with Mexico, Singapore and Viet Nam submitting their tariffs in the HS 2012 nomenclature. These economies' preferential tariffs needed to be aligned to HS 2017 before any tariff analysis could be performed in the report. Canada's 2018 preferential tariffs, while submitted in HS 2017, also required tariff line transposition to be aligned with its 2021 MFN rates. This entire process done manually took WTO statisticians approximately three months to align tariffs to a common nomenclature before they could be processed for analysis. In contrast, this new pilot methodology allows each economy's tariffs data to be concorded automatically in less than five minutes. Human verification can then follow, taking about half a day.

Faced with the inherent challenges of statistical work on tariffs and trade data, this pilot initiative is motivated by a need to augment the productivity of existing staffing resources and introduce efficiency gains to traditional tariff processing methodologies, primarily through the application of computational and automation techniques from Natural Language Processing (NLP), using both artificial intelligence (AI) and lexical similarity methods.

As of the time of this writing, automated tariff line level HS transposition using NLP techniques applied to WTO transposition methodology has not yet been documented in mainstream international trade literature. This paper details the requirements, process and logic behind this pilot process.

III. LITERATURE REVIEW

Natural Language Processing (NLP) is one among many sub-branches of artificial intelligence (AI) that specifically deals with text as data. It covers a range of computational methods including **text similarity**, which is defined in simple terms as a process of measuring the similarity between text strings. Comparisons can be done using **lexical similarity**, which measures the closeness of word sets of same or different languages, or **semantic similarity**, which creates a quantitative measure of the likeness of *meaning* between words or phrases, such as through the use of text embeddings.⁷ Lexical similarity does not typically use AI and instead relies on rules-based, statistical or algorithmic methods. In contrast, semantic similarity uses a range of AI techniques such as embedding models (e.g. Word2Vec or Global Vectors for Word Representation (GloVe)), or deep learning models (e.g. Siamese neural networks or recurrent neural networks (RNNs)). **Transformer models** are an advanced evolution of both deep learning and embedding models that create *contextual* embeddings capturing semantic meaning based on Google's novel **self-attention mechanism** architecture, which focuses on different parts of the input sequence to capture relationships between words in a context-dependent way ([Vaswani, et al., 2017](#)).

Lexical similarity models are usually evaluated using metrics such as cosine or Jaccard similarity applied to vectorized text data such as Bag of Words (BoW), Term Frequency Inverse Document Frequency (TF-IDF) or Positive Pointwise Mutual Information (PPMI) vectors. Semantic similarity techniques use embeddings generated by models such as Word2Vec or Sentence-BERT (SBERT) ([Tripathy, 2024](#)). These models are employed in semantic textual similarity (STS) tasks, which apply cosine similarity to semantically meaningful sentence embeddings generated from large text datasets ([Reimers and Gurevych, 2019](#)). Compared to traditional string distance measures, these approaches are better suited for tasks that require deep, context-based semantic understanding, and are most commonly applied in semantic search, question answering, and text clustering.

Depending on the requirements of the corpus⁸, lexical models may provide straightforward methods for calculating text similarity without requiring model training or large computational resources ([van der Loo, 2014](#)). However, lexical models are limited to surface-level matching and may not capture different words with similar meanings. On the other hand, semantic models can understand the intent behind queries, categorize texts based on meaning, and can handle synonyms and varied

⁷ **Embeddings** in NLP broadly refer to vector representations of the meaning of words. TF-IDF vectors for representing words are also called embeddings, although the term is sometimes more strictly applied only to dense vectors like Word2Vec ([Jurafsky and Martin, 2024](#)).

⁸ A **corpus** (plural corpora) in NLP refers to a computer-readable, organized collection of text or audio data that often include a wide range of documents, texts or voices in one or more languages ([Jurafsky and Martin, 2024](#)).

expressions more effectively; however, they require sophisticated modelling and have much higher computational requirements.

TF-IDF and cosine similarity as a straightforward lexical similarity metric

Cosine similarity is one of the most used lexical string distance metrics. It is commonly applied to document similarity and clustering, ranking search results (information retrieval), recommendation systems, and topic modelling. It measures the similarity between two vectors of an inner product space through the cosine of the angle between them, in effect determining if they are pointing in roughly the same direction ([Data Mining \(Third Edition\), 2012](#)). A cosine similarity of 1 denotes two identical vectors, while 0 denotes completely opposite vectors.

The cosine similarity is defined mathematically as:

$$\text{sim}(\mathbf{x}, \mathbf{y}) = \cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

where $\mathbf{x} \cdot \mathbf{y}$ is the dot product of the vectors and $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ indicate their magnitudes (Euclidean norms). Compared to other common methods such as Jaccard which measures similarity based on the absence or presence of terms, cosine similarity has the advantage of also considering the frequency of terms. This can be beneficial if certain keywords or phrases and their frequency are more critical when matching product descriptions.

Cosine similarity is effective for use with vectorized text such as TF-IDF matrices.

Term Frequency Inverse Document Frequency (TF-IDF) combines the frequency of occurrence of terms with their importance across the corpus. Term frequency (TF) is computed as the ratio of term occurrence over the total number of terms, while inverse document frequency (IDF) is calculated by penalizing terms that occur more frequently, calculated as the logarithm of the ratio of the total number of documents to the number of documents containing the term.

$$TF(t, d) = \frac{\text{number of times } t \text{ appears in } d}{\text{total number of terms in } d}$$

$$IDF(t) = \log \frac{N}{1 + df}$$

$$TF - IDF(t, d) = TF(t, d) * IDF(t)$$

Terms with higher TF-IDF scores thus significantly represent the content of the document ([Tripathy, 2024](#)).

In the case of product description matching, the main requirement is keyword matching across vectors of reasonable length, and for this reason a computationally efficient combination of TF-IDF vectorization and cosine similarity is a reasonable baseline method for automated HS transposition.

Sentence-BERT as a context-based semantic similarity model

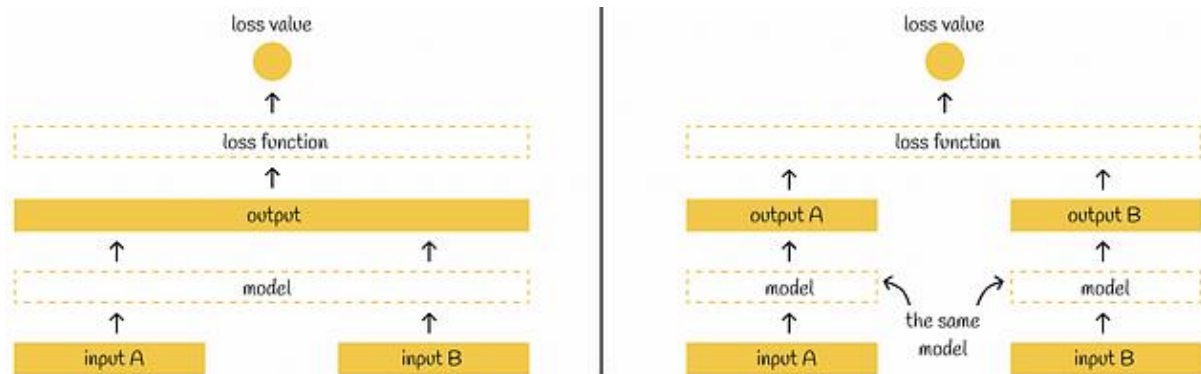
Sentence-BERT (SBERT) is an extension of the original Bidirectional Encoder Representation from Transformers (BERT) model introduced by Google in 2018. Compared to BERT, which produces embeddings for individual words, SBERT is designed to create fixed-size embeddings for entire sentences, making it easier to compare and match them. SBERT works by fine-tuning the BERT model to generate meaningful sentence embeddings. It uses a Siamese network architecture where two BERT models with shared weights process pairs of sentences. The main difference is that the BERT model accepts two inputs at the same time, while the SBERT model accepts two inputs in parallel so that they are not dependent on each other (See **Figure 1**).

During training, SBERT learns to produce embeddings such that semantically similar sentences have closer embeddings in the vector space. This enables SBERT to effectively capture the overall meaning of sentences, making it suitable for tasks like semantic textual similarity, clustering, and information retrieval by comparing these embeddings using measures like cosine similarity. Computationally, SBERT is also more efficient than BERT because it reduces the need to re-process sentences for

every comparison. It achieves this by generating fixed-size sentence embeddings once and then using simple vector operations for comparisons.

As a comparative model to measure string matching performance, a pre-trained SBERT model is chosen as a second method to perform automated HS transposition.

Figure 1: Non-Siamese (BERT) architecture on the left compared to Siamese (SBERT) architecture on the right.



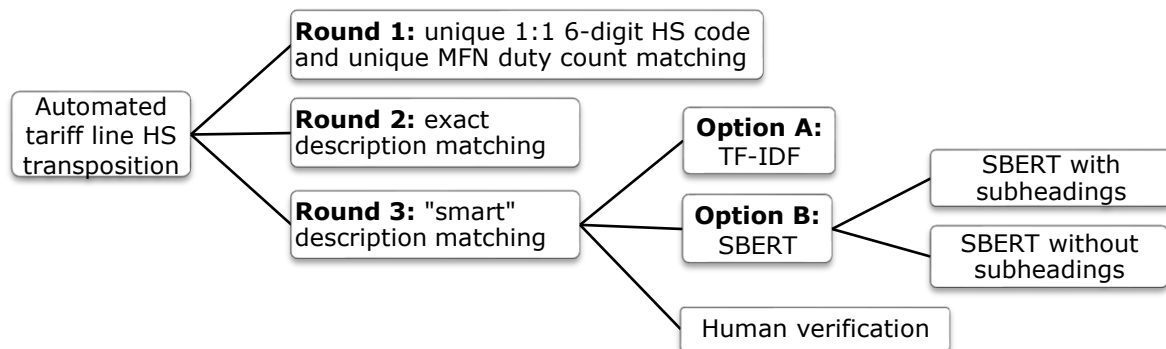
Source: (Efimov, 2023).

IV. METHODOLOGY

A. Overview of the process

Figure 2 outlines the end-to-end process for automated HS transposition, following three staged rounds of concordance (further detailed in **Section C**):

Figure 2: Diagram of the automated tariff line HS transposition methodology



Round 1 follows a unique 1:1 6-digit HS code and unique MFN duty count matching, applicable only to cases where no other matching options exist for the tariff line level HS code. **Round 2** follows exact product description matching across all remaining unmatched tariffs as the next best method after Round 1 mechanics have been exhausted. These matches are done using the combined tariff line, 6-digit level and 4-digit level product descriptions to enrich the text string. For **Round 3** "smart" description matching, two methods are compared to perform automated tariff line HS transposition using cosine similarity. **Option A** uses TF-IDF vectorization and **Option B** uses embeddings generated from a pre-trained SBERT model. An SBERT version **with subheadings** is also compared to an SBERT version **without subheadings**.

To account for the wide variation in tariff complexity across economies, *two tariff scenarios* were considered for a transposition between HS 2017 and HS 2022: **Economy A**⁹ exhibited a "simple" tariff scenario with standardized product descriptions, while **Economy B** showed a "complex" tariff scenario with more highly specific product descriptions with inline references to regulations, quotas, and technical as well as foreign language terms. To evaluate the usefulness of subheadings in the string matching, two versions of SBERT were used for Economy B's tariffs, one including subheadings and one without. All results were then evaluated by a manual human line-by-line verification.

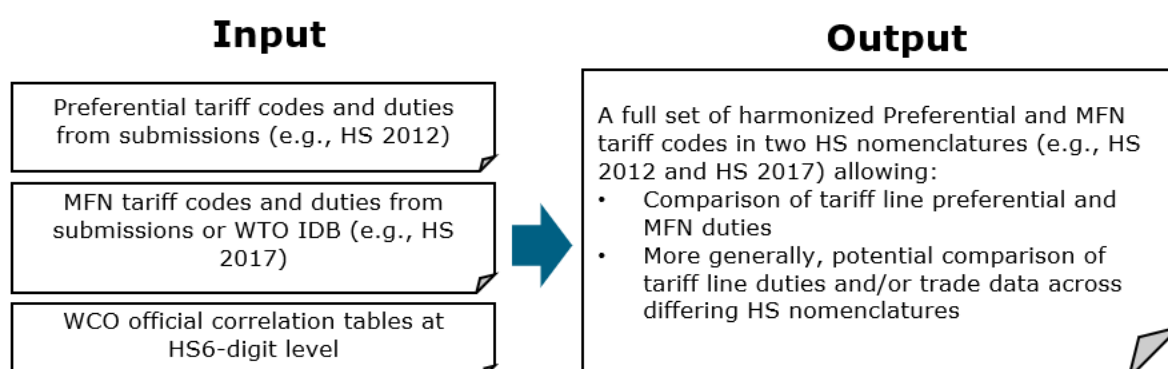
How is cosine similarity used in Round 3 "smart" description matching?

The cosine similarity scores generated by either the TF-IDF or the SBERT method are combined with the WTO standard HS transposition methodology (utilizing WCO-issued HS 6-digit correlation tables) to perform a "smart" string matching of HS product descriptions at the tariff-line level, by choosing the tariff code with the highest cosine similarity from the concorded pool of tariff codes at the 6-digit level.

Importantly for Round 3, the cosine similarity matching is only run *within the smaller pool of possible tariff lines*, thus significantly improving the chances of a successful match.

Figure 3 below details the input requirements of the automated HS transposition process to produce the final output file. The output file can then be directly used to compare previously disjoint data such as tariff line level preferential and MFN duties, or more generally tariff line level data across different HS nomenclatures.

Figure 3: Inputs and outputs to the automated HS transposition process



Notably for the output file, the entire Preferential tariff codes list is retained, with each HS code matched to a single MFN duty. Any additional MFN tariff codes and duties that are not used are excluded from the final output.

B. Pre-processing of tariff files

The tariff line level preferential and MFN tariff codes to be matched are first pre-processed and standardized. Sources for preferential tariffs are usually Member submissions¹⁰ to the WTO as part of Transparency Mechanism obligations for trade agreements. Sources for MFN tariffs are either Member submissions or the WTO IDB database. The submitted preferential tariffs excel file should ideally contain data columns for "TL", "TLS", "Description", "Duty" and "Quota":

- The *Tariff line (TL) column* is the minimum required data for the HS transposition to be successful.
- If not provided, a *Tariff line suffix (TLS) or "ex-out" column* (for products that only contain part of the referenced subheading as specified in the description) is generated automatically. This is important to avoid duplicate entries within the same tariff line.

⁹ Economy names are anonymized.

¹⁰ Or non-Member submissions, for trade agreements between WTO Members and other economies.

- The *Description column* should be provided at tariff line level. While tariff line descriptions greatly augment the string matching, the process will default to product matches generated from a combination of the 6 and 4-digit descriptions in the rare cases where it is not provided.
- The *Duty column* should be provided at tariff line level and is often compared with the matched MFN duty.
- A *Quota column* is required, but only for agreements that have in-quota and out-quota tariff conditions for some products.
- A general *Note or Remarks column* is also often included to specify quotas, ex-outs, or other category information that is not provided in the other columns.

The MFN tariffs file should include "TL", "TLS", "Description", "Duty", as well as "Quota" columns that will be matched to the corresponding preferential tariff code. The Description column should be provided at tariff line level. If the data is retrieved from the WTO IDB database all of this information is fully available.

The two data files do not need to have the same number of tariff codes, however all the tariff codes should be present in their respective tariff nomenclature at the HS 6-digit level, satisfying standard WTO tariff completeness checks.

Standardization includes providing full descriptions pulled from the tariff line, 6-digit and 4-digit levels. The descriptions for 6 and 4-digit level codes are pulled from the Basnomen reference file prepared by the WTO Integrated Database (IDB) after each new WCO revision of the HS nomenclature. This master file contains a full transposition across all HS versions and subheading levels down to the harmonized 6-digit level, including product descriptions in English, French and Spanish. These descriptions are parsed to both the tariff line level preferential tariffs descriptions provided by the economy and the tariff line MFN tariffs descriptions processed by the WTO IDB.

Utilizing the full product descriptions allows the highest probability of a successful string match. Notably, this avoids issues where the tariff line description contains only generic text such as "other", "not elsewhere specified", or else too concise descriptions such as "virgin" [coconut oil] that may lead to multiple matches from disparate headings or subheadings. An "HS6 count" column is added to indicate the number of 6-level lines for each MFN tariff. This is important for generating the first round of string matches.

For the TF-IDF lexical similarity baseline method, the full product descriptions followed standard pre-processing by stemming, converting all characters to lowercase, and removing punctuations, numbers, stopwords, and extra whitespaces to retain the words with the most relevance. For the SBERT model, the only pre-processing done was to convert to lowercase and remove whitespaces to retain context from the full product description. Subheadings, where available, were added only during Round 3 (cosine similarity round). The tm() and SnowballC() packages in R allow these text mining transformations ([Feinerer, 2024](#) and [Bouchet-Valat, M., 2023](#)).¹¹ For the TF-IDF method, the strings were tokenized and vectorized by creating a Document-Term Matrix (DTM) to determine the frequency counts of terms in the strings, and finally transformed into a TF-IDF matrix that adjusts the word counts by their importance, all using the word2vec() package in R ([Wijffels, J., et al., 2023](#)).¹² For the SBERT model, the reticulate() package ([Kalinowski, T., Ushey, K., & Allaire, J., 2024](#)) was used to create a Python interface in R and run the sentence-transformers module ([Reimers, N., & Gurevych, I., 2023](#)). Finally, the vectors and embeddings generated from the TF-IDF method and the SBERT model were compared using cosine similarity from the stringdist() package ([van der Loo, M. \(2014\)](#)).

¹¹ String pre-processing notably improves string matching for minor discrepancies between descriptions that are otherwise identical.

¹² Tokenizing and vectorizing observably improves string matching for descriptions that are more challenging to classify and correctly assigns them to the "Other" or "not elsewhere specified" category within the WCO-concordance pool of possible matches.

C. HS transposition

The actual transposition process starts with an initially blank MFN duty rate column that is filled up over several rounds:

1. **Round 1** concords unique 1:1 6-digit HS code and unique MFN duty count matches, in the particular case of when there is either no change in the 6-digit HS code (typically around 98% of 1:1 matches) or there is a change in the 6-digit HS code to one new line (the remaining 2%), *while having only one possible tariff line level MFN duty* that can be matched to that tariff line. This is the priority round for matching tariff lines, as the tariff line simply copies the corresponding MFN rate since there is only one possible match. Notably, these cases also correspond to the "one-to-one relationships" and "merging" categories defined by standard WTO transposition methodology, with the added condition of *only having one possible tariff line MFN duty*:

Table 1: Cases covered by Round 1 HS transposition

Category	Description	Example
one-to-one relationships	One 6-digit subheading corresponds exactly to one 6-digit subheading	3402 Organic surface-active agents [...] No change HS 2017 3402.90 - Other HS 2022 3402.90 - Other Change to one new line HS 2017 3402.12 -- Cationic HS 2022 3402.41 -- Cationic
merging	Two or more 6-digit subheadings into one 6-digit subheading	9114 Other clock or watch parts. Original lines HS 2017 9114.10 - Springs, including hair-springs 9114.90 - Other Merged to one line HS 2022 9114.90 - Other

Source: Transposition of Members' CTS files to the HS 2022 nomenclature ([WTO, 2023](#))

2. **Round 2** concords exact full description matches, combining the tariff line, HS6 and HS4 level descriptions without referencing any HS codes.
3. **Round 3** combines the WCO HS 6-digit concordance with cosine similarity matching run on the generated TF-IDF vectors and SBERT embeddings to retrieve the best match. The product description with the highest cosine similarity score is chosen (e.g. closest to a value of 1).

Round 2 and Round 3 cases correspond to the "splitting" and "complex cases" categories defined by standard WTO transposition methodology:

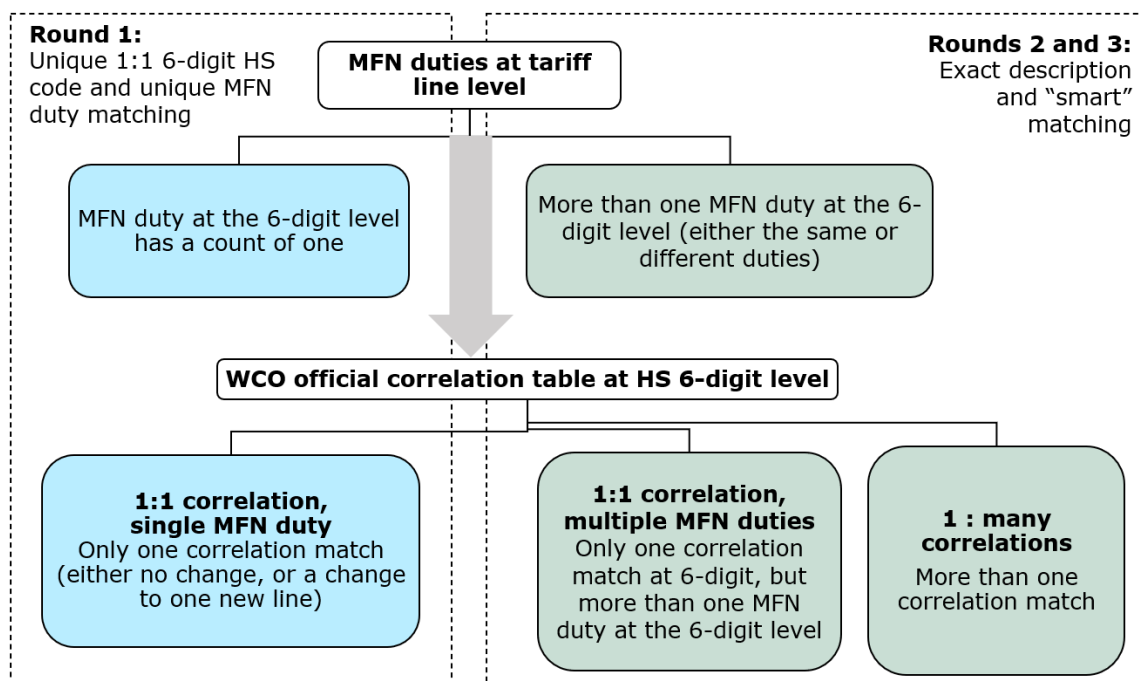
Table 2: Cases covered by Round 2 and Round 3 HS transposition

Category	Description	Example
splitting	One 6-digit subheading into two or more 6-digit subheadings	Original line HS 2017 1510.00 Other oils and their fractions, obtained solely from olives, whether or not refined, but not chemically modified, including blends of these oils or fractions with oils or fractions of heading 15.09. Splitting to two new lines

		<p>HS 2022 1510.10 - Crude olive pomace oil 1510.90 - Other</p>
complex cases	Involving both splitting and merging of whole or part of different 6-digit subheadings	<p>6116 Gloves, mittens and mitts, knitted or crocheted.</p> <p>Original lines HS 2017 6116.10 - Impregnated, coated or covered with plastics or rubber 6116.9 - Other: 6116.91 - - Of wool or fine animal hair 6116.92 - - Of cotton 6116.93 - - Of synthetic fibres 6116.99 - - Of other textile materials</p> <p>Splitting and merging at the same time HS 2022 6116.10 - Impregnated, coated, covered or laminated with plastics or rubber 6116.9 - Other: 6116.91 - - Of wool or fine animal hair 6116.92 - - Of cotton 6116.93 - - Of synthetic fibres 6116.99 - - Of other textile materials</p> <p>Remark: 6116.91 is split into two subheadings, and at the same time, both 6116.10 and part of 6116.91, 6116.92, 6116.93 and 6116.99 are merged into 6116.10</p>

Source: Transposition of Members' CTS files to the HS 2022 nomenclature ([WTO, 2023](#))

Figure 4: Diagram depicting tariff sets built for Rounds 1, 2, and 3



How are the tariff sets for Rounds 1, 2 and 3 built?

- a. A new "MFN count at HS6" column is added to the MFN file, counting the number of tariff-line level MFN duties for each HS 6-digit.
- b. The "MFN count at HS6" column is joined with the WCO official correlation table using the HS 6-digit code, allowing for the identification of cases where there is only one possible MFN duty associated with each HS 6-digit code.
- c. The updated WCO official correlation table is split into three sets: a **1:1 correlation, single rate** set which goes to Round 1; a **1:1 correlation, multiple rates** set and a **1: many correlations** set, both of which go to Rounds 2 and 3.

For Round 3 matches, additional reference columns retrieved from the smaller pool of possible tariff lines are also computed: the HS 6-digit level tariff rate average and HS 6-digit level maximum possible tariff rate are useful benchmarks for additional verification of the automated tariff matching; if identical to the duty rate allocated to the concorded tariff line, there is no need for additional checking.

D. Note on subheadings

Certain economies contain additional subheadings at the 8-digit level subheading or beyond in addition to the harmonized 4- and 6-digit level to clarify some products. This may pose problems for classification if the tariff line product description retained for string matching only contains 4- and 6-digit level descriptions that omit this additional level of detail.¹³ Likewise, unlike 6-digit level subheadings, not all tariff line level product descriptions may have an equivalent 8-digit level subheading, and the level of subheadings may also vary depending on the HS nomenclature. For instance, Economy B's HS 2022 nomenclature, provided at the 9-digit level tariff line, only has 8-digit level subheadings for 8.2% of tariff lines but has 6-digit level subheadings for 17.7% of all tariff lines. Also, its HS 2017 nomenclature only contains up to the 6-digit level subheading.

The inclusion of HS 2022 8-digit subheadings in the full product descriptions for automated HS transposition between HS 2017 and HS 2022 may have mixed effects. In some cases, it may improve the quality of the string match, but in others it may have the potential to introduce confounding effects on the matching algorithm depending on the complexity and semantic distance of the additional text from the original tariff line description.

E. Results evaluation

In the absence of a national correlation table, evaluating HS transposition results typically requires manual human verification, as automated methods alone cannot ensure accurate assessment.

Rounds 1 and 2 represent "priority" HS transposition matches, consisting of non-AI-based results that are deemed accurate and do not require further human verification.

The primary conservative benchmark for performance in Round 3 will be the **minimum accuracy**, defined as the percentage of instances where both TF-IDF and SBERT identify *the same* valid tariff match as a human statistician. This metric provides a baseline, while the **actual accuracy** for either TF-IDF or SBERT may be higher when considering the percentage of *all cases* where either method successfully matched tariffs.

Subsequently, the **overall accuracy** will be calculated as the total percentage of accurate matches across Rounds 1, 2, and 3, based on the minimum accuracy alone. This conservative measure will facilitate a clear comparison of TF-IDF and SBERT models' effectiveness in HS transposition.

¹³ For instance, in Economy B's HS 2022 nomenclature, the 6-digit level subheading **030399**: "Other" is further clarified by the two 8-digit level subheadings, which in turn have 9-digit tariff line products also with identical descriptions "Other".

030399 --Other [Fish parts]:

03039991 ----(1) *Nishin (Clupea spp.)*, *Tara (Gadus spp., Theragra spp. and Merluccius spp.)*, *Buri (Seriola spp.)*, *Saba (Scomber spp.)*, *Iwashi (Etrumeus spp., Sardinops spp. and Engraulis spp.)*, *Aji (Trachurus spp. and Decapterus spp.)* and *Samma (Cololabis spp.)*

030399919 ----Other

03039999 ----(2) Other [referring to "Other fish other than Nishin et. al."]

030399999 ----Other

V. RESULTS

After running both transposition methods, the final output is a transposed tariffs table ("TC" table, in WTO IDB parlance) providing the full set of correlated HS codes in two HS nomenclatures, with each preferential tariff line allocated a corresponding MFN duty code.

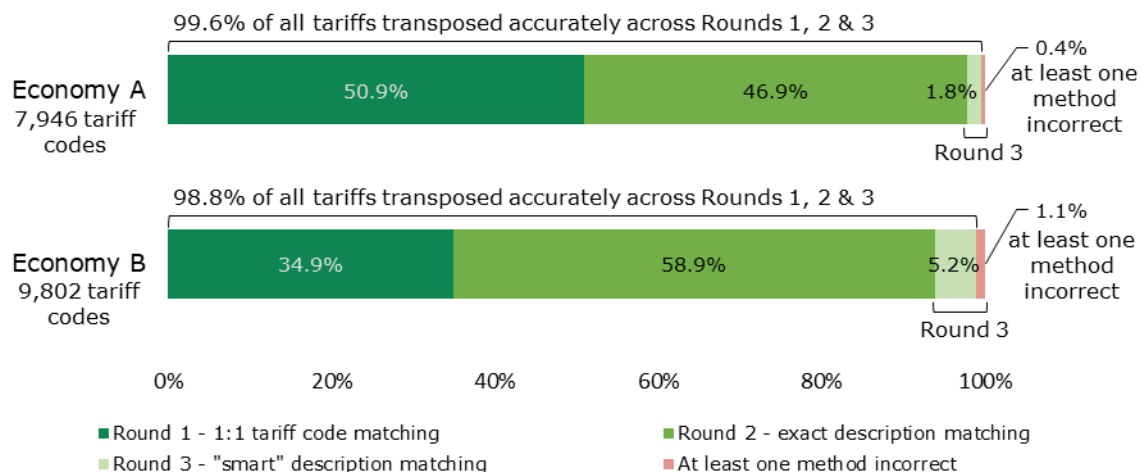
This section provides main and detailed results of automated HS transposition from HS 2017 to HS 2022: 1) a simple case for Economy A's¹⁴ tariffs comparing TF-IDF and SBERT methods; and 2) a complex case for Economy B's tariffs, which compares TF-IDF and SBERT methods with and without the inclusion of subheadings.

Hypothetically, SBERT would perform better than TF-IDF, and SBERT with subheadings would perform better than SBERT without subheadings. To test out these hypotheses, a full line-by-line human verification check was done to verify the results.

1. Overview of main results

Chart 1 shows the breakdown of results from the automated HS transposition, in terms of accuracy as well as the three staged rounds of concordance: Round 1 – 1:1 tariff code matching, Round 2 – exact description matching and Round 3 – "smart" description matching.

Chart 1: Breakdown of automated HS transposition results, by processing round and accuracy



The overall accuracy of automated HS transposition across Rounds 1, 2 and 3 is 99.6% for Economy A and 98.8% for Economy B (without subheadings). Notably, the majority of accurately performed automated tariff code matching (97.8% for Economy A and 93.8% for Economy B) is already covered between Rounds 1 and 2. This underscores the result that novel automation techniques relying on tariff characteristics at the tariff line level paired with exact description matching already provide significant improvements to traditional HS transposition methodologies even without the assistance of AI techniques.

A further 1.8% (for Economy A) and 5.2% (for Economy B¹⁵) of accurately performed automated tariff code matching are then covered by Round 3 by the TF-IDF and SBERT (AI-based) "smart" description matching. Finally, the remaining 0.4% Economy A and 1.1% Economy B tariff codes—equivalent to 33 of 7,946 tariff codes and 102 of 9,802 tariff codes respectively—are the lines where at least one method of either TF-IDF or SBERT incorrectly matched the product description.

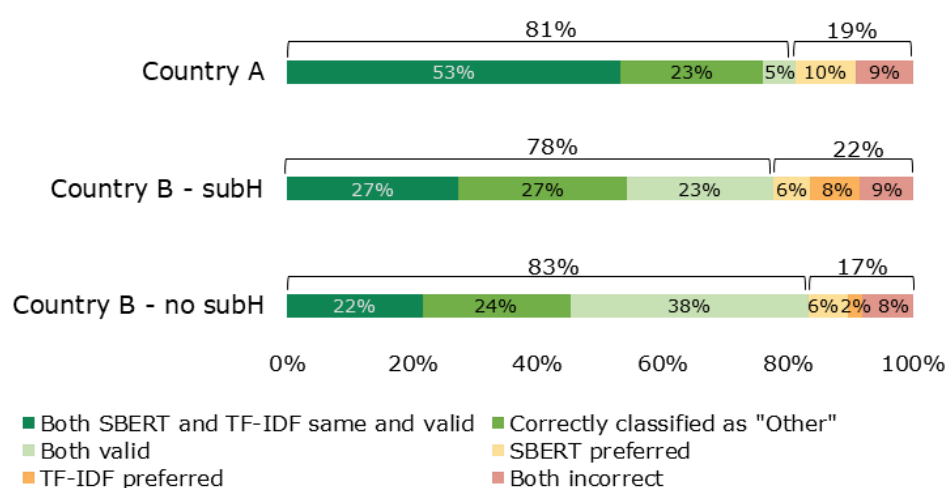
¹⁴ Economy names are anonymized.

¹⁵ Based on SBERT without subheadings; for SBERT with subheadings accurate matches comprise 4.8%, while inaccurate matches comprise 1.4%.

Notably, while AI-based automated tariff transposition covered by Round 3 accounts for a smaller percentage of tariff codes (2.2% of Economy A tariff codes and 6.2% of Economy B tariff codes) compared to Rounds 1 and 2 across the entire process, the human equivalent of this "smart" description matching is also the subset of manual tariff transposition that requires the most work for trade statisticians. For example, the number of lines captured in Round 3 (175 lines for Economy A and 611 lines for Economy B) would typically require around a month to manually concord given the complexity of tariff nomenclatures.

Chart 2 below breaks down the Round 3 results comparison of SBERT and TF-IDF automated HS transposition for a) Economy A, b) Economy B – with subheadings, and c) Economy B – without subheadings, in terms of categories and accuracy.

Chart 2: Comparison of Round 3 results by scenario, category and accuracy



Of the three, Economy B without subheadings (no subH) and Economy A yielded the highest minimum accuracy (81% and 83% respectively), while Economy B with subheadings (subH) performed the lowest (78%). SBERT performed the best in Economy A's scenario (91%) and the worst in Economy B with subheadings (84%), which was the only scenario where it was outperformed by TF-IDF (86%). On the other hand, at least one of the methods was incorrect in 17–22% of all automated HS transposition matches, and *both* methods were incorrect in 8–9% of all matches. While indeed the range of results may vary based on the overall complexity of a economy's tariff regime, these scenarios provide an indication of results that may be encountered across other economies' tariff profiles.

The rest of this section provides detailed breakdowns and analyses of each category of results across both simple and complex scenarios.

2. Detailed results for a simple scenario

The table below details the results for the simple scenario, in decreasing order of confidence from Round 1 to Round 3:

Table 3: Simple scenario – Economy A HS 2017 Preferential tariffs with HS 2022 MFN tariffs

When Concordinated	Count	Percentage
Round 1 – 1:1 tariff code matching	4,044	50.9%
Round 2 – exact description	3,727	46.9%
Round 3 – "smart" description matching	175	2.2%

Round 3 TF-IDF and SBERT transposition results for Economy A's tariffs were then individually validated by human line-by-line HS transposition and tagged for the following categories:

Table 4: Summary of Round 3 results for 175 Economy A automatically correlated tariffs:

Results	Human Verification	Count	Description
142 correct matches (81.1%)	Both SBERT and TF-IDF same and valid	93	Both SBERT and TF-IDF correctly classified the tariff line to the same HS product code.
	Correctly classified as "Other"	40	Both SBERT and TF-IDF correctly classified the tariff line to the same HS product code with the description "Other".
	Both valid	9	The original product description was ambiguous; both TF-IDF and SBERT chose equally valid product codes.
33 matches where one or both were incorrect (18.9%)	SBERT preferred	17	The SBERT model chose the more accurate product code.
	TF-IDF preferred	0	The TF-IDF method chose the more accurate product code.
	Both incorrect	16	Neither TF-IDF nor SBERT chose the human-verified priority product code, for various reasons.
TOTAL		175	Out of 7,946 national tariff lines, there were 175 HS product lines left over from Round 1 and 2 that required string matching.

In the breakdown above, using either TF-IDF or SBERT to automate HS transposition for a simple scenario such as Economy A's tariffs yields a minimum of 81.1% accuracy. In the case of SBERT, the total accuracy was 90.1% after factoring in the lines where it performed better than TF-IDF. For TF-IDF, the accuracy of 81.1% was retained since it did not perform better than SBERT for any tariff lines.

In 93 cases, both SBERT and TF-IDF chose the same valid HS code. In most of the cases, the matched product description was usually easy to classify by reading the text and contained similar wording to the original. Economy A's tariffs pose a simple, typical tariff scenario where most product descriptions are largely similar and standardized across tariff nomenclatures.

In 40 cases, both SBERT and TF-IDF correctly classified the tariff line to the catch-all product description "Other". This code is the same one that a human verifier would choose when none of the other codes accurately capture the product to be classified. This is also one of the most common scenarios for HS transposition.

In 9 cases, both SBERT and TF-IDF chose valid but different codes. Examples of these tariffs are detailed below:

Table 5: Examples of cases where SBERT and TF-IDF both chose valid but different tariffs¹⁶

Original HS code	Description	Method	Matched code	Description
293971000	Cocaine, ecgonine, levometamfetamine, metamfetamine (INN), metamfetamine racemate; salts, esters and other derivatives thereof	SBERT	293972000	Cocaine, ecgonine; salts, esters and other derivatives thereof
		TF-IDF	293945000	Levometamfetamine, metamfetamine (INN), metamfetamine racemate and their salts
36030000	Safety fuses; detonating fuses; percussion caps; detonatings caps;	SBERT	36032000	Detonating cords
		TF-IDF	36031000	Safety fuses

¹⁶ Notably, Round 3 results are limited to the "best choice" matches chosen by the TF-IDF and SBERT algorithms, and will ultimately require final human verification, for instance in the case of exact splits where the new tariff lines are all equally valid, and the ideal human-based solution would be to generate tariff line suffix (TLS) ex outs for each new split.

	igniters; electric detonators			
890391000	Sailboats, with or without auxiliary motor	SBERT	890323000	[Sailboats] Of a length exceeding 24 m
		TF-IDF	890321000	[Sailboats] Of a length not exceeding 7.5 m
97060000	Antiques of an age exceeding one hundred years	SBERT	97061000	[Antiques of an age exceeding one hundred years] Of an age exceeding 250 years
		TF-IDF	97069000	Other [Antiques of an age exceeding one hundred years]

In 17 cases, SBERT chose a more accurate match over TF-IDF, examples of which are shown in the table below:

Table 6: Examples of cases where SBERT outperformed TF-IDF

Original HS code	Description	Method	Matched code	Description
15091000	Virgin [olive oil]	SBERT	15093000	Virgin olive oil ¹⁷
		TF-IDF	15094000	Other virgin olive oils
22087060	[Liqueurs] ¹⁸ containing more than 9% vol., but not more than 14 % volper l	SBERT	22087060	[Liqueurs] containing more than 9% vol., but not more than 14 % volper l
		TF-IDF	22087020	[Liqueurs] Containing not more than 1.15 % vol
24022090	[Cigarettes] Not exceeding in weight 0.8 kg actual tobacco content per 1,000 cigarettes	SBERT	24022090	[Cigarettes] Not exceeding in weight 0.8 kg actual tobacco content per 1,000 cigarettes
		TF-IDF	24022010	[Cigarettes] Exceeding in weight 0.8 kg actual tobacco content per 1,000 cigarettes
38220020	Other [Diagnostic or laboratory reagents]	SBERT	38229000	Other [Diagnostic or laboratory reagents]
		TF-IDF	38221900	Other [Diagnostic or laboratory reagents] [For malaria]
94051001	[Chandeliers] Of plastics	SBERT	94051901	[Chandeliers] Of plastics
		TF-IDF	94051910	Other [Chandeliers]

In the majority of cases, SBERT was able to pick out key wording or more precise phrasing that was missed by TF-IDF. For instance, in the case of **24022090** it was able to precisely match the phrase *"Not exceeding in weight 0.8 kg"* compared to TF-IDF, which matched with the phrase *"Exceeding in weight 0.8 kg"*.¹⁹ Likewise as in the case of **94051001**, TFIDF may automatically choose the definitive "Other" catch-all category and miss a more precise category option.²⁰

¹⁷ It could be argued that both choices are valid. However the simple literal interpretation points to **15093000**.

¹⁸ Brackets indicate summarized descriptions parsed from higher-level subheadings.

¹⁹ Examining the raw matched data shows that the preprocessing for TF-IDF indeed removed both the negation as well as the numeric values from the original phrase, retaining only: *"exceeding weight kg actual cigars cheroots cigarillos cigarettes tobacco tobacco substitutes cigarettes containing tobacco"* in the MFN rate.

²⁰ Retrieving the raw string used for the TF-IDF matching shows that the adjective "Other" is indeed removed from the MFN rate: *"luminaires lighting fittings including searchlights spotlights parts thereof elsewhere specified included illuminated signs illuminated name plates like permanently fixed light source parts thereof elsewhere specified included"*.

How important is pre-processing in string matching?

These examples highlight that pre-processing plays a key role in the string matching performance of both SBERT and TF-IDF. On the one hand, pre-processing may remove "noise" to improve the quality of string matching. However it may also remove key contextual information from the inclusion of punctuations, numeric values, prepositions, clarificatory adjectives and adverbs (particularly "Not" and "Other", which are particularly important in tariff product descriptions). In this respect, SBERT's superior context-based string matching provides an advantage over TF-IDF.

Notably, for simple cases such as Economy A's tariffs, there was not much difference in the performance of SBERT compared to TF-IDF. However it is worthwhile to note that SBERT was able to make superior tariff matches for 17 product lines compared to TF-IDF, while TF-IDF did not perform better than SBERT in any tariff matches.

Finally, in 16 cases, both SBERT and TF-IDF chose an incorrect product match. In some cases the correct category for "Other" could not be determined as there were multiple "Other" categories at various subheading levels which could not all be captured in the string matching. The same situation was also true for other generic descriptions such as "Each ply exceeding 6 mm in thickness", which had multiple exact matches across different subheading levels. In other cases there was too much complex wording in the description, such as "**95089011**: Roundabouts, swings, shooting galleries and other fairground amusements", that was classified as "Other" by both methods rather than under "Fairground amusements." In one case there was a typo ("ceramics" vs "ceramincs") in the description.

3. Detailed results for a complex scenario

Notably, Economy B's tariff descriptions are among the most complex across all WTO submissions. They may include highly specific language linking back to laws, quota or other agreement provisions well as very detailed product categories, in addition to subheadings at the 5- and 7-digit level that could not be captured in the string matching.²¹

The summary table of results below reflects this complexity, with many more unmatched tariff lines after Rounds 1 and 2 than for Economy A.

Table 7: Complex scenario – Economy B HS 2017 Preferential tariffs with HS 2022 MFN tariffs

When Concorded	Count	Percentage
Round 1 – 1:1 tariff code matching	3,418	34.9%
Round 2 – exact description	5,773	58.9%
Round 3 – "smart" description matching	611	6.2%

To determine the effect of subheadings inclusion in the string matching, two versions of Round 3 transposition results for Economy B's tariffs were run (with and without 8-digit level subheadings), and are summarized below, along with a comparison of results:

²¹ For example, see HS product code **100390011**: [Other] [Barley] [For feeding purposes] "Imported by Economy B Government according to Article 42 of "Act on Stabilization of Supply, Demand and Prices of Staple Food", imported to be purchased and sold by Economy B Government in response to a joint application by seller to and purchaser from Economy B Government according to Article 43 of the law or imported with certification of Minister of Agriculture, Forestry and Fishery according to the Cabinet Order concerning wheat and others provided by the Cabinet Order provided in column 3 of paragraph 1 of Article 45 of the law"

Table 8: Summary of Round 3 results for Economy B's 611 automatically correlated tariffs, with subheadings compared to without subheadings

With subheadings		
Results	Human Verification	Count
474 correct matches (77.6%)	Both SBERT and TF-IDF same and valid	167
	Correctly classified as "Other"	164
	Both valid	143
137 matches where one or both were incorrect (22.4%)	SBERT preferred; match subheadings	5
	SBERT preferred	31
	TF-IDF preferred	48
	Both incorrect	53
TOTAL		611

Without subheadings		
Results	Human Verification	Count
509 correct matches (83.3%)	Both SBERT and TF-IDF same and valid	132
	Correctly classified as "Other"	144
	Both valid	233
102 matches where one or both were incorrect (16.7%)	SBERT preferred	38
	TF-IDF preferred	14
	Both incorrect	50
TOTAL		611

The more complex scenario of Economy B tariffs resulted in a minimum accuracy of 77.6% for all cases, increasing to 83.5% for SBERT with subheadings and 85.4% for TF-IDF. Surprisingly, the TF-IDF method performed slightly better than SBERT. In contrast, removing subheadings from the SBERT model raised the minimum accuracy to 83.3% for all cases, increasing to 89.5% for SBERT and 85.6% for TF-IDF.

In the simple scenario case, SBERT's total accuracy was 90.1% after factoring in the lines where it performed better than TF-IDF. For TF-IDF, the accuracy of 81.1% was retained since it did not perform better than SBERT for any tariff lines.

In the complex scenario, SBERT's highest total accuracy was 89.5% in the version without subheadings, while TF-IDF's performance was 85.6%. When subheadings were included in SBERT's tariff matching, its accuracy went down to only 83.5%, which was lower than TF-IDF's (net) performance of 85.4%.

3.1 Both correct or valid matches

The count of automatically correlated tariffs classified under the categories where both TF-IDF and SBERT chose *the same correct tariff match* ("Both SBERT and TF-IDF same and valid" and "Correctly classified as 'Other' ") were slightly higher with the inclusion of subheadings in the SBERT model, totalling 331 (54.2% of all cases) compared to 276 (45.2% of all cases) in the SBERT model without subheadings.

However, the count of correlated tariffs under the category "Both valid" – where TF-IDF and SBERT chose different but equally valid tariff matches – nearly doubled, rising from 143 to 233 cases with the exclusion of subheadings. This increase represents a shift from 23.4% to 38.1% of all cases. Notably, most of these cases involved highly complex tariff descriptions where the original product description was sufficiently ambiguous as to allow multiple options to be equally valid.²²

The table below provides some examples of these cases:

Table 9: Examples of cases where TF-IDF and SBERT tariff matches were both valid

²² For instance, **040150111**: [Milk and cream] [Of a fat content, by weight, exceeding 10 %] For "the Pooled Quota of other milk products" or in some cases where the tariff-line description was left blank, **271012120**: [Light oils and preparations] NA.

Original HS code	Description	Method	Matched code	Description
040291121	[Milk and cream] Concerning milk and cream in this subheading 1-(2) and 2, for the quantity (quota) stipulated by a Cabinet Order on the basis of 1,500 ton ...	SBERT-subH	040291110	Whipped cream in pressurized containers
		SBERT-no subH	040291110	Whipped cream in pressurized containers
		TF-IDF	040291290	Other [Milk and cream]
210690291 ²³	[Other] [Food preparations not elsewhere specified or included] NA	SBERT-subH	210690518	[Other] [Food preparations not elsewhere specified or included]
		SBERT- no subH	210690119	[Other] [Food preparations not elsewhere specified or included] Other
		TF-IDF	210690247	(b) [Other] [Food preparations not elsewhere specified or included] Other
292090010 ²⁴	Carbonic esters and their derivatives	SBERT-subH	292090019	Other [Esters...]: Carbonic esters and their derivatives
		SBERT- no subH	292090011	Other [Esters...]
		TF-IDF	292090090	Other [Esters...]

For instance, **040291121** broadly refers to milk and cream imports under a specified volume-based quota, and was matched to both valid categories of "Whipped cream" and "Other". In some cases such as **210690291**, the exclusion of subheadings forced tariff matches between descriptions with identical "Other" text. Finally, in certain cases such as **292090010**, the product description was sufficiently specified, however there were blanket categories specified by "invisible" subheadings that allowed for multiple tariff matches to be valid. In such cases and other similar situations with highly ambiguous product descriptions, tariff matches were counted as valid for both TF-IDF and SBERT models.

3.2 One or both incorrect matches

On the other hand, comparing the tariff matches where either one or both TF-IDF and SBERT matches were incorrect highlighted some differences in performance for each method.

With subheadings, TF-IDF performed better than SBERT for 48 tariff line matches, however upon removing subheadings most of these matches moved to either "both SBERT and TF-IDF same and valid" or "both valid" categories, ostensibly due to the improved performance from SBERT catching up with TF-IDF. In the version without subheadings, only 14 tariff matches were left where TF-IDF was still more accurate than SBERT.

The table below shows cases of TF-IDF performing better than SBERT and SBERT-no subH:

Table 10: Examples of cases where TF-IDF outperformed SBERT

²³ Tariff-line product description is blank. **210690518** is the definitive "Other" category for the full 6-digit. **210690119** does not capture the 8-digit subheading **210690111**: *Of a milkfat content, by weight, not exceeding 30 %: QV* ; and **210690247** does not capture the tariff-line level subheading **210690240**: "Konnyaku".

²⁴ **292090011** and **292090090** do not capture the 8-digit subheading **29209001**: "Carbonic esters and their derivatives", but are covered by it.

Original HS code	Description	Method	Matched code	Description
030559090	[Dried fish] Samma (Cololabis spp.)	SBERT-subH	030559020	[Dried fish] (1) Nishin (Clupea spp.), Buri (Seriola spp.), Saba (Scomber spp.) and red-eye round herring (Etrumeus spp.)
		SBERT- no subH	030559020	[Dried fish] (1) Nishin (Clupea spp.), Buri (Seriola spp.), Saba (Scomber spp.) and red-eye round herring (Etrumeus spp.)
		TF-IDF	030559090	[Dried fish] Other
441899291	Cross Laminated Timber (CLT) product that has ... a width of not less than 300mm and a length of not less than 900mm, and is attached by technical materials including a grading label of Economy B Agricultural Standard (JAS) ...	SBERT-subH	441882110	[Cross-laminated timber (CLT or X-lam)]: Of a length not less than 900 mm and of a width not less than 300 mm
		SBERT- no subH	441882190	Other [Cross-laminated timber (CLT or X-lam)]
		TF-IDF	441882110	[Cross-laminated timber (CLT or X-lam)] Of a length not less than 900 mm and of a width not less than 300 mm
290339011	Methyl bromide	SBERT-subH	290369010	Brominated derivatives
		SBERT- no subH	290369010	Brominated derivatives
		TF-IDF	290361000	Methyl bromide (bromomethane)

For instance, in the case of **030559090**, TF-IDF more accurately matched Samma to the "Other" [dried fish] category, rather than matching it incorrectly to another breed of fish, as done by both SBERT models. Failing to match on context or exact wording, they seem to have matched only the closest available text, namely "spp." In the complexly worded case of **441899291**, TF-IDF and SBERT with subheadings more accurately matched the exact wording "Of a length not less than 900 mm and of a width not less than 300 mm" compared to SBERT without subheadings, which failed to highlight this key phrase and instead matched to the catch-all "Other" category. Finally, in the case of **290339011**, despite having a very short product description, both SBERT models failed to identify the technical key phrase "Methyl bromide" and instead matched with "Brominated derivatives" instead of an exact match, which TF-IDF was able to accomplish.

When does TF-IDF hold an advantage over SBERT?

Notably, in cases with long and highly complex product descriptions, particularly those involving scientific, technical or foreign language terms, TF-IDF seemed to have an advantage from its lexical similarity approach of measuring only string distance rather than relying on context. For a majority of cases, this allowed it to more accurately identify the definitive catch-all "Other" category (usually with the highest value product code). On the other hand, the highly mathematical parsing structure of tariff product descriptions that combine tariff-line, subheading, 6-digit and 4-digit descriptions also may not provide the ideal narrative context for SBERT's semantic model, which typically relies on context gleaned from grammatically cohesive paragraphs of dense and longer text.

Lastly, from **Table 8** previously, SBERT with subheadings outperformed TF-IDF for 36 tariff matches, 5 of which were facilitated by exact matches from the inclusion of subheadings. Without subheadings, this total count increased slightly to 38, broken down thusly with some newly incorrect tariff matches being compensated for by newly corrected tariff matches:

Table 11: Summary of changes to results between SBERT-subH and SBERT-no subH

SBERT-subH category	SBERT-no subH category	Count	Result
---------------------	------------------------	-------	--------

SBERT preferred; match subheadings	Both incorrect	5	Results changed to "both incorrect"; the missing subheadings contained the text for exact matches.
SBERT	SBERT	31	Results remained the same; SBERT-no subH still chose the correct tariff match over TF-IDF.
Both incorrect	SBERT	7	Results changed from "both incorrect" to SBERT; tariff matches were facilitated by excluding subheadings
TOTAL		38	Total tariff lines in SBERT-no subH where SBERT outperformed TF-IDF.

This illustrates that *removing* subheadings from SBERT does not significantly impact the total counts of tariff matching for SBERT-preferred matches. Instead, most of its advantage comes from "upgrading" previously incorrect SBERT matches to the "both correct or valid" category.

When it does affect SBERT matches, the results may be mixed. In some cases such as **390110020**, the removal of subheadings for SBERT may degrade the quality of the tariff matching by removing important keywords. In other cases such as **100590099**, removing subheadings may also improve the quality of the tariff matching when the product descriptions being matched are sufficiently complex. The table below illustrates these examples.

Table 12: Examples of comparative cases between SBERT-subH and SBERT-no subH

Original HS code	Description	Method	Matched code	Description	Human verify
390110020	Linear low density polyethylene	SBERT-subH	390110029	Linear low density polyethylene	SBERT-subH
		SBERT-no subH	390110061	Other [Bio-polyethylene]	
		TF-IDF	390110061	Other [Bio-polyethylene]	
100590099	Less than 10% by weight passes through a woven metal wire cloth sieve with an aperture of 13.2mm Treatment for Peru	SBERT-subH	100590010	Other [Maize] for feeding purposes Note: The imports under this item are to be used as materials for fodder and feed under the supervision of the Customs	SBERT-no subH
		SBERT-no subH	100590099	Other [Maize]	
		TF-IDF	100590020	Other [Maize] Popcorn	

Do subheadings provide a net benefit to SBERT?

Overall, excluding subheadings led to a net improvement in SBERT's accuracy, indicating that in this scenario subheadings more frequently introduce confounding elements into the string matching process rather than facilitate it. Further testing on other datasets is recommended.

4. Computational efficiency

The generation of embeddings may also be resource-intensive in terms of time and computational power. The table below shows the amount of time taken to run each of the two methods discussed:

Table 13: Comparison of time taken for SBERT and TF-IDF:

Method	Average time taken (minutes)	
	Economy A (simple scenario)	Economy B (complex scenario)
SBERT	4.4	6.1
TF-IDF	1.2	1.6

The time taken for Economy B (a complex tariff scenario) is longer than for Economy A (a simple tariff scenario), taking 1.2 times as long for SBERT and 1.3 times as long for TF-IDF. This could be due to the more complex structure of Economy B's tariff descriptions. Additionally, SBERT code took around 4 times longer than TF-IDF to run.

This potentially has implications when working with Big Data but is relatively less of a concern for tariff datasets.

VI. TARIFF VERIFICATION

Both the simple and complex tariff scenarios discussed above indicate that a level of human verification is still ultimately required to ensure the quality of results after the automated tariff transposition, depending on the quality of the original tariffs datasets. This section highlights different strategies for optimizing this process:

1. HS 6-digit MFN average tariff as a benchmark for accuracy

The *primary goal* of tariff line level HS transposition is to match an MFN tariff to each product line. Thus, the most expedient way to avoid the need for tedious tariff verifications is to disregard products where all possible matching options have identical MFN tariffs. While the pure exercise of matching a correct tariff line remains relevant, the correct MFN tariff will be reflected regardless of the transposition results.

For instance in the case of Economy A, 106 of 175 correlated tariff lines had identical MFN tariffs for all possible transposition matches. Disregarding these lines meant that only **69 lines** needed to be further verified for HS transposition matching. For Economy B, there were **450 lines** out of 611 total tariff lines that required further verification based on this criterion.

What tariff lines require final manual verification?

With the end goal of matching the appropriate MFN tariff to each tariff line, out of a set of approximately 10,000 tariff-line level HS codes, only the remaining tariff lines that do not match their nearest 6-digit MFN average tariff require further verification. For Economy A this was 69 tariff lines or 0.9% of 7,946 total lines, while for Economy B it was 450 tariff lines or 4.6% of 9,802 lines.

2. Wisdom of the crowds as a benchmark for accuracy

The summary table below breaks down how often the two methods were able to match the correct (or incorrect) tariff line, either together or separately:

Table 14: Breakdown of HS tariff transposition results where SBERT and TF-IDF chose the same product codes.

Economy A		
Accuracy	Both chose same	Both different
Both correct or valid	133	9
Only one correct	-	17
Both incorrect	1	15
Economy B – with subheadings		
Accuracy	Both chose same	Both different
Both correct or valid	331	143
Only one correct	-	84
Both incorrect	12	41
Economy B – without subheadings		
Accuracy	Both chose same	Both different
Both correct or valid	276	233
Only one correct	-	52
Both incorrect	10	40

The predominant trend appears to be that **correct matches occur most frequently when both methods choose the same HS code**, with 133 cases (76.0%) for Economy A's simple tariffs scenario, 331 cases (54.2%) for Economy B's complex tariffs using SBERT with subheadings, and 276 cases (45.2%) for Economy B's complex tariff without subheadings.

As tariffs become more complex, the number of cases where both methods chose different but valid tariff matches also increases, from 9 cases (5.0%) for Economy A to 143 cases (23.4%) and 233 cases (38.1%) for Economy B.

Finally, the number of cases where both methods chose incorrect tariff matches hardly changed between the two scenarios of Economy B with and without SBERT subheadings. Instead, the removal of subheadings and subsequent improved SBERT performance migrated many tariffs from TF-IDF having the advantage of "only one correct" to "both valid" matches.

These results highlight that both TF-IDF and SBERT methods matching to the same tariff line is often a good indication that an accurate match has been made.

3. Cosine similarity scores as a benchmark for accuracy

The cosine similarity score for incorrect SBERT matches ranged from 0.45 to 0.98, with an average score of 0.81, while for TF-IDF matches it ranged from 0.08 to 0.99 with an average score of 0.80. Meanwhile for correct matches SBERT cosine similarities ranged from 0.19 to 1.00 with an average of 0.82, and for TF-IDF it ranged from 0.19 to 1.00 with an average score of 0.80. In general, SBERT gave higher overall cosine similarity scores than TF-IDF, and there were both correct and incorrect matches made at all levels of cosine similarity scores for both methods. This makes a cosine similarity threshold score a less reliable benchmark for determining matching accuracy.

VII. CONCLUSIONS

The study highlights that incorporating NLP techniques into HS transposition processes offers substantial potential to enhance the efficiency and accuracy of tariff analysis, making it an invaluable tool for trade statisticians. However, it is essential to maintain human oversight to ensure quality and reliability within AI-assisted HS transposition workflows, as highlighted in Round 3 results.

This paper details a novel approach to automating a traditionally manual, labour-intensive and error-prone process in international trade statistics. It highlights some key features comparing the suitability of TF-IDF vectorization and SBERT embeddings methods to enable "smart" tariff

description matching. It also discusses results in the specific context of the WCO HS nomenclature and standard WTO methodologies for tariff processing.

Key findings include the following:

1. Automated HS transposition using NLP techniques can significantly augment traditionally manual HS transposition by reducing the processing time for manual HS transposition from approximately 2 to 3 weeks to potentially half a day (up to 30 times faster), with accuracy rates of 99.6% for simpler tariff scenarios and 98.8% for complex ones out of a standard set of 10,000 tariff codes.

This study examined the HS 2017 to HS 2022 transposition for Economy A and Economy B, noting that results may vary depending on the specific complexity of each economy's tariff regime. Nonetheless, the findings suggest that automated HS transposition could greatly support statisticians, allowing them to focus on the simpler task of verifying the final output while reducing the risk of manual errors.

2. Non-AI based automation techniques covered by Rounds 1 and 2—relying on tariff characteristics at the tariff line level paired with exact description matching—already provide significant improvements to traditional HS transposition methodologies.

Notably, the majority of accurately performed automated tariff code matching (97.8% for Economy A and 93.8% for Economy B) is already covered between Rounds 1 and 2. This underscores the result that (non-AI-based) automation techniques relying on tariff characteristics at the tariff line level paired with exact description matching already provide significant improvements to traditional HS transposition methodologies. A further 1.8% (for Economy A) and 5.2% (for Economy B²⁵) of accurately performed automated tariff code matching are then covered by Round 3 by the TF-IDF and SBERT (AI-based) "smart" description matching.

3. While AI-based automated HS transposition covered by Round 3 accounts for a smaller percentage of all tariff codes (2.2% of Economy A tariff codes and 6.2% of Economy B tariff codes), the human equivalent of this "smart" description matching is the subset of manual HS transposition that requires the most work.

The majority of manual work required for traditional HS transposition entails the laborious line-by-line comparison of product descriptions across all possible tariff line descriptions within the corresponding 6-digit subheading(s). For example, the number of lines captured in Round 3 (175 lines for Economy A and 611 lines for Economy B) would typically require around a month to manually concord given the complexity of tariff nomenclatures.

4. At the minimum, approximately 80% of all Round 3 "smart" description matches are correct.

The minimum accuracy, defined as cases where *both* TF-IDF and SBERT chose the same valid tariff match, was similar for both simple and complex tariff scenarios. The simple scenario had an 81.1% minimum accuracy for simple cases and the complex scenario had an 83.3% minimum accuracy. This indicates that at the minimum, approximately 80% of all "smart" description matches in the pilot scenarios were correct.

As mentioned previously, the range of results may vary based on the overall complexity of a economy's tariff regime. Notably, for complex scenarios where a large number of product descriptions are *sufficiently ambiguous* for the purposes of interpretation, there may be a large number of similarly valid tariff match options, which can result in a higher minimum accuracy. In the majority of cases, the tariff rates for such options are also identical.

5. Final verification of automated HS transposition results means covering only the subset of Round 3 results where tariff lines do not match their nearest 6-digit MFN rate average; these account for around 1-5% of all tariff codes.

²⁵ Based on SBERT without subheadings; for SBERT with subheadings accurate matches comprise 4.8%, while inaccurate matches comprise 1.4%.

The goal of HS transposition is to match the appropriate most-favoured-nation (MFN) rate to each tariff line. After the automated HS transposition, only the remaining tariff codes that do not match their nearest 6-digit MFN rate average require further final verification. For the simple scenario of Economy A this was 69 tariff codes (0.9% of 7,946 tariff codes), while for the complex scenario of Economy B it was 450 tariff codes (4.6% of 9,802 lines).

6. In general, SBERT (AI-based HS transposition) performs better than TF-IDF, with a total accuracy of around 90% compared to around 80 to 85% for TF-IDF in the pilot scenarios. For this study, the inclusion of subheadings does not provide a net benefit to SBERT performance. Further testing is recommended.

In the simple scenario case, SBERT's total accuracy was 90.1% after factoring in the codes where it performed better than TF-IDF. For TF-IDF, the accuracy of 81.1% was retained since it did not perform better than SBERT for any tariff codes. In the complex scenario, SBERT's highest total accuracy was 89.5% in the version without subheadings, while TF-IDF's performance was 85.6%. When subheadings were included in SBERT's tariff matching, its accuracy went down to only 83.5%, which was lower than TF-IDF's (net) performance of 85.4%.

Although subheadings may occasionally aid in the tariff matching of certain codes, their overall impact in the tested scenario tended to reduce SBERT's performance by introducing potentially confounding text elements during string matching. Their inclusion should be tested further across more tariff datasets.

7. The particular syntax and specificity of tariff product descriptions impacts the performance of TF-IDF and SBERT, and in certain cases of highly technical tariffs TF-IDF showed advantages due to its straightforward matching algorithm.

Two factors specific to tariff product descriptions are particularly relevant. First, the structural parsing of tariff descriptions is often highly mathematical, potentially preventing it from achieving optimal performance compared to grammatically cohesive, dense paragraph text. Second, long and complex tariff descriptions involving scientific, technical or foreign language terms may not assimilate easily into SBERT's general language training corpus.

Hence, while SBERT maintains an overall performance advantage for the majority of tariff matches, in a few cases TF-IDF's use of straightforward numeric string distance rather than context for string matching may have certain advantages, particularly for highly complex tariff descriptions where the best solution is often the (simply worded) definitive catch-all "Other" category.

The end result is a fully functional end-to-end automated HS transposition process requiring minimal human verification as compared to traditional processes.

VIII. FURTHER WORK

In future, further refinement and testing on diverse datasets is recommended to optimize these methods for broader application. For instance, other models such as OpenAI's Text Embedding model could be compared with SBERT to compare performance rates.

In applying these methods to other workstreams, tariff line analysis comparing MFN duties with bound duties, preferential duties, or MFN duties in another nomenclature could see significant improvements. Additionally, tariffs and trade analysis typically conducted at the 6-digit level could be extended to the more detailed tariff line level, facilitating the development of richer to enable more in-depth analyses. For example:

- In Trade Policy Reviews (TPRs), the comparison of the MFN applied duty with the bound duty at the tariff line level is required to identify possible breaching (e.g. MFN rate > Bound rate), since both lists are often expressed in different HS nomenclatures.

- The United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP) has developed a [Trade Intelligence and Negotiation Adviser \(TINA\)](#) tool that provides negotiating advice on tariffs with comparative advantages at the HS 6-digit level; however economies often negotiate tariffs at very detailed tariff line or "ex-out" levels that go beyond the aggregated HS 6-digit level.
- Generally, economic research involving trade agreements may compare MFN, preferential, or bound tariffs as well as trade data across different time periods; however harmonized time series data is often only available at the HS 6-digit level.

IX. BIBLIOGRAPHY

1. Bouchet-Valat, M. (25 April 2023). *Package 'SnowballC'*. Comprehensive R Archive Network. Available at: [SnowballC: Snowball Stemmers Based on the C 'libstemmer' UTF-8 Library \(r-project.org\)](#)
2. Efimov, D. (12 September 2023). *Large Language Models: SBERT — Sentence-BERT*. Medium blog article. Available at: [Large Language Models: SBERT — Sentence-BERT | by Vyacheslav Efimov | Towards Data Science](#)
3. Feinerer, I., Hornik, K., & Meyer, D. (2008). *Text Mining Infrastructure in R: tm Package*. Journal of Statistical Software, 25(5), 1-54. Available at: <https://doi.org/10.18637/jss.v025.i05>
4. Han, Kamber and Pei (2012). *Data Mining (Third Edition). Cosine Similarity in Text Mining Applications*. Morgan Kaufmann Publishers. Available at: [Data Mining. Concepts and Techniques, 3rd Edition \(The Morgan Kaufmann Series in Data Management Systems\) \(sabanciuniv.edu\)](#)

5. Jurafsky, D. and Martin, J. (2024). *Speech and Language Processing*. Stanford University. Available at: [ed3bookfeb3_2024.pdf \(stanford.edu\)](#)
6. Kalinowski, T., Ushey, K., & Allaire, J. (September 5 2024). *Package 'reticulate'*. R-project. Available at: [reticulate: Interface to 'Python' \(r-project.org\)](#)
7. Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks*. Technische Universitat Darmstadt. Available at: [SBERT Official Site](#)
8. Tripathy, S. (5 January 2024). *Exploring Contextual Text Similarity: A Dive into Machine Learning Techniques*. Medium blog article. Available at: [Exploring Contextual Text Similarity: A Dive into Machine Learning Techniques | by Swarup Tripathy | Medium](#)
9. United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP). (2024). *Trade Intelligence and Negotiation Adviser (TINA)*. Available at: [TINA: Trade Intelligence and Negotiation Adviser](#)
10. van der Loo, M. (2014). *The Stringdist Package for Approximate String Matching*. The R Journal, 6(1), 111-122. Available at: [loo.pdf \(r-project.org\)](#)
11. Vaswani, A., Shazeer, N., Parmar, N., et al., (2017). *Attention Is All You Need*. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. Available at: <https://doi.org/10.48550/arXiv.1706.03762>
12. Wijffels, J., et al. (7 October 2023). *Package 'word2vec'*. Comprehensive R Archive Network. Available at: [word2vec: Distributed Representations of Words \(r-project.org\)](#)
13. World Customs Organization (WCO). (2018). *The Harmonized System: A Universal Language for International Trade 30 Years On* (pp. 19-20). Available at: [World Customs Organization \(wcoomd.org\)](#)
14. World Customs Organization (WCO). (2024). *What is the Harmonized System (HS)?*. Available at: [World Customs Organization \(wcoomd.org\)](#)
15. World Trade Organization (WTO). (27 February 2023). *Factual Presentation on the "Comprehensive and Progressive Agreement for Trans-Pacific Partnership Appendix 8 Peru*. Document ID: WT/REG395/1/Add.8. Available at: [docs.wto.org](#)
16. World Trade Organization (WTO). (28 April 2023). *Harmonized System 2022: Transposition methodology for Members' CTS files*. Committee on Market Access. Document ID: G/MA/410. Available at: [docs.wto.org](#)
17. World Trade Organization (WTO). (31 May 1995). *Minutes of the Meeting of 25 April 1995*. Committee on Market Access. Document ID: G/MA/M/1. Available at: [docs.wto.org](#)