KYG Trade™
Know Your Goods



WHITE PAPER: CASE STUDY

Optimizing Global Tariff Classification with KYG Trade's ChatHTS AI Assistant (Kay)

Validated Case Study on Accuracy and Operational Efficiency

Abstract

This case study summarizes a proof-of-concept (POC) evaluation of KYG Trade's generative AI assistant, Kay (ChatHTS), for global tariff classification. The system was used to classify a sample of 121 industrial products into the US, Mexico, EU and China under Harmonized System (HS) codes.

Using production-representative data from a global industrial manufacturer, the POC produced measurable outcomes, including a 98% accuracy rate (after excluding items with insufficient data) and an overall efficiency improvement described as ~75% across the 121-item sample. The POC materials also report a 74% reduction in hours and costs.

These results illustrate how generative AI can help trade compliance organizations improve classification consistency and reduce routine manual effort while preserving appropriate human oversight for exceptions.

Executive Summary

- **Accuracy:** 98% accuracy rate after excluding items with insufficient data (15 items excluded).
- **Benchmarking:** Results were compared to the client's reference HS codes and a Big Four consulting firm's trade classification database.
- **Automation and effort reduction:** 74% of items were classified correctly without a human in the loop (HIL); overall efficiency improvement described as ~75% across the sample.
- **Operational impact:** The POC materials report a 74% reduction in hours and costs.
- **Scope:** Proof-of-concept evaluation using a 121-item sample of industrial products.

Methodology and Definitions

KYG Trade's ChatHTS assistant (Kay) was used to classify a 121-item sample of industrial products. Results were evaluated by comparing outputs against the client's reference HS codes and a Big Four consulting firm's trade classification database. Items identified as having insufficient data were excluded from the accuracy calculation (net excluded: 15).

Key Definitions

- **HS codes:** Harmonized System (HS) commodity classification codes used to determine global tariff treatment.

- **POC:** Proof-of-concept evaluation performed to measure outcomes on a defined sample before broader deployment.
- **Human-in-the-loop (HIL):** A workflow in which a human reviewer validates or dispositions exceptions or lower-confidence outputs.
- **OneShot:** The initial single-pass classification interaction used in the POC workflow (i.e., the first determination and question set captured before any re-initiation or follow-up).
- **Windchill:** A product lifecycle management (PLM) system used to manage and provide access to product engineering information such as part records, bills of materials, drawings, and related attributes.
- **Accuracy rate (post-exclusion):** The accuracy rate reported after removing items excluded for insufficient data (98% after excluding 15 items).
- **Excluded for insufficient data:** Items for which available product information was not sufficient to support a reliable classification outcome within the POC (net excluded: 15).
- **Automation rate:** Share of items classified correctly without a human in the loop (74% in this POC sample).

Summary of Key Findings

- **Accuracy (post-exclusion):** After exclusions, the measured accuracy rate was 98%.
- **High automation rate:** 74% of items were classified correctly without human-in-the-loop (HIL) involvement.
- **Efficiency and cost outcomes:** Overall efficiency improvement of ~75% and a 74% reduction in hours and costs.

Introduction

Tariff classification is a high-impact control within global trade compliance programs, influencing landed cost, preferential treatment, duty recovery, and regulatory exposure. Multinational enterprises often face variability in classification outcomes due to distributed teams, incomplete product data, and differing interpretations across jurisdictions.

This document presents an anonymized POC case study with a global industrial manufacturer, evaluating KYG Trade's ChatHTS assistant (Kay) for HS classification performance and operational efficiency. The objective is to provide trade compliance leaders with transparent, quantified outcomes and clear conditions under which the results should be interpreted.

Use Case Overview

Objectives

The POC evaluated ChatHTS performance in three areas:

- Assessing AI-assisted classification of industrial products under HS codes.
- Reducing routine human effort through automation while retaining human-in-the-loop (HIL) review for exceptions.
- Quantifying time and cost impacts across the sample.

Process

- **Dataset:** 121 industrial products provided by the industrial manufacturer.
- **Comparison:** Classification results were compared against the client's reference HS codes and matches in a Big Four consulting firm's trade classification database.
- **Metrics Evaluated:** Accuracy, degree of human-in-the-loop (HIL) involvement, and operational savings.

Key Findings

- **Automation:** 74% of items were classified correctly without a human in the loop (HIL).
- **Overall efficiency:** The POC materials describe an overall efficiency improvement of ~75% across the 121-item sample.
- **Hours and cost impact:** The POC materials report a 74% reduction in hours and costs.
- **Risk Reduction:** Misclassification risk materially reduced through AI-driven consistency.
- **Tangible Cost Reductions:** Demonstrated tangible cost reductions through automation.

Category	Description	Count
Excluded from Test Universe	Excluded by client – insufficient data	12
	Excluded by KYG – insufficient data	7
	Add back 4 of 12 by KYG	-4
	NET excluded for insufficient data	15
False Negative	Favors KYG (see note)	41
True Negatives	Incorrect Classification	2
True Positives	Match Big Four reference database	63
	Grand Total	121

Accuracy is calculated on the in-scope population after excluding items with insufficient data. The numerator is 104, comprised of 63 “True Positives” (matches to the Big Four reference database) plus 41 “False Negatives” determined to favor KYG. The denominator is 106, comprised of the 121 grand total minus 15 net excluded for insufficient data. This yields 104/106 ≈ 98%.

NOTE: Two of the client “insufficient data” results were converted to “Matched” after re-initiating OneShot.

Business Impact

Before KYG Implementation

- Manual processes prone to inconsistencies.
- High dependency on experts for error resolution.
- Misclassifications caused regulatory and financial setbacks.

After KYG Implementation

- Improved classification accuracy and consistency through AI-driven insights.
- Substantial efficiency and cost savings with reduced human involvement.
- Future scalability ensured by KYG’s integration capabilities and continuous improvement focus.

Conclusion

The POC results indicate that KYG Trade’s ChatHTS AI assistant can support global tariff classification with high measured accuracy (98% after excluding items with insufficient data) and reduced manual effort. For trade compliance leaders, these outcomes suggest a path to improve consistency, standardize decision support, and allocate specialist capacity toward exception handling and governance.

For additional detail on the approach, controls, and implementation considerations, contact KYG Trade.

Exhibit A: Analysis of False Negatives (n=41)

Research Question	KYG Findings
Is Control reference HS a valid code?	3 invalid codes + 1 EU v US mismatch
Was the client reference HS code present?	Only 1 had the client reference HS code shared. Therefore, KYG was not relying on prior reference codes.
Was Item Number in client knowledge base?	35 item numbers were not found in the client knowledge base
Was KATID completed?	40 items were not completed by client human testers, of which 13 only had initial OneShot.
Was COI set correctly?	1 item was set up with the wrong country of import; thus HS code did not match
Was the OneShot description correct?	9 OneShot product descriptions needed to be corrected. Once corrected, HS matched.
Would humans find answers in public domain?	6 items' product attributes were not found in an internet search and therefore there wasn't sufficient information
Would humans find answers in product attributes?	18 items did not have required data in product attributes and therefore there wasn't sufficient data
Would humans find answers in Windchill data?	8 did not have answers in the Windchill data
Would humans find answers in drawings and images?	None of the 121 items in the sample had image files.
Did the client knowledge base have a relevant product with correct HS code?	13 did not have valid and/or relevant HS codes in the client knowledge base
Was HS run after the client knowledge base was implemented?	34 of the classifications were run prior to the October 9th client knowledge base update.
Did Kay recommend the same, or better HS code?	41 had the same or better HS code

Note: Exhibit A reproduces the false-negative disposition analysis from the POC readout materials, presented here in text format for ease of review.

Confidential — Unauthorized use and disclosure prohibited. © 2026 KYG Trade, Inc. All rights reserved.