Modelo multidimensional para la gestión de proyectos de traducción en la era de la inteligencia artificial: flujos de trabajo, calidad, riesgo terminológico e intervención humana

Gabriela A. González
Cuestiones del Lenguaje, Universidad Nacional de Mar del Plata
Traductora freelance
traductoragonzalez@gmail.com

Pedro L. Luchini
Universidad Nacional de Mar del Plata
luchinipedroluis@gmail.com

Resumen

La Gestión de proyectos de traducción atraviesa una transformación acelerada impulsada por flujos de trabajo híbridos y por la adopción creciente de inteligencia artificial. Este artículo propone un modelo multidimensional que integra análisis teórico, comparación de modelos operativos y estudios de caso ilustrativos construidos a partir de patrones reales observados en agencias de traducción pequeñas y medianas. Se examinan comparativamente tres enfoques de producción —tradicional, híbrido y centrado en la IA— y sus implicancias para la calidad, los costos, la congruencia terminológica y la toma de decisiones operativas. Los casos ilustrativos permiten visualizar cómo la integración de traducción automática y modelos generativos afecta la gestión terminológica, el control de calidad multicapa y la sostenibilidad profesional, destacando riesgos como el *drift* terminológico, las alucinaciones semánticas y la sobreconfianza en outputs automatizados. Metodológicamente, el estudio combina revisión documental, observación indirecta y análisis comparativo para fundamentar un marco que articula supervisión humana estratégica, gobernanza del flujo y gobernanza terminológica en entornos multilingües mediados por IA. Finalmente, se discuten las limitaciones del enfoque conceptual y se proponen líneas futuras de investigación vinculadas con métricas de calidad, validación empírica y prácticas responsables de integración tecnológica, junto con recomendaciones operativas para agencias, traductores y clientes.

Palabras clave: gestión de proyectos de traducción; inteligencia artificial; flujos de trabajo híbridos; gobernanza terminológica; calidad multicapa

1. Introducción

El crecimiento exponencial del contenido digital y la expansión a mercados globales han consolidado a los servicios multilingües como un componente crítico de la comunicación internacional. La demanda actual exige velocidad, escalabilidad y adecuación cultural, transformando la Gestión de proyectos de traducción en una disciplina estratégica que integra conocimiento lingüístico, tecnológico y operativo. Entendida como el conjunto de procesos mediante los cuales se planifica, organiza, ejecuta y controla un encargo multilingüe, esta gestión coordina recursos humanos, tecnológicos y operativos para cumplir con estándares de calidad, plazos y requerimientos del cliente. Para ello, se apoya en el diseño de flujos de trabajo, que en este contexto se definen como arquitecturas operativas que secuencian tareas, roles, herramientas y puntos de control de calidad desde la recepción del encargo hasta la entrega final. En consecuencia, el gestor de proyectos ya no es un mero coordinador, sino un profesional capaz de diseñar flujos, evaluar riesgos y asegurar coherencia estilística en un entorno donde la presión por entregas rápidas es constante.

En este escenario, se consolidan flujos híbridos (combinan automatización y supervisión humana) que integran traducción humana, posedición (revisión y corrección humana de una salida automática), motores de traducción automática neuronal (NMT, por sus siglas en inglés: *Neural Machine Translation*), modelos de lenguaje a gran escala (LLMs, por sus siglas en inglés: *Large Language Models*), y controles automatizados de calidad (QA, por sus siglas en inglés: *Quality Assurance*). Este ecosistema exige alfabetización tecnológica, análisis crítico y gobernanza del flujo (reglas, roles y puntos de control que organizan la producción de punta a punta), además de gestión terminológica y comunicación intercultural. Paralelamente, el *AI-shift* (reconfiguración de procesos, roles y expectativas por la adopción masiva de inteligencia artificial) redefine expectativas de clientes, estructuras tarifarias y criterios de calidad, e introduce riesgos como el drift terminológico —esto es, la deriva gradual e inconsistente en el uso de términos clave dentro de un mismo proyecto o entre iteraciones, que erosiona la coherencia con glosarios y memorias— y las alucinaciones semánticas —salidas plausibles pero incorrectas que incorporan o alteran significado

respecto del texto origen (p. ej., añaden información, cambian matices o reformulan relaciones)—, además de la sobreconfianza en modelos generativos.

El objetivo de este capítulo es proponer un modelo multidimensional para comprender y gestionar esta transformación, integrando teoría, análisis comparativo y estudios de caso ilustrativos basados en patrones reales. Se comparan tres modelos operativos — tradicional, híbrido y centrado en la IA— y se analizan sus implicancias para la calidad, los costos, la consistencia terminológica y la sostenibilidad profesional, destacando la necesidad de control de calidad multicapa y gobernanza terminológica. Finalmente, se organiza la exposición en torno al marco teórico, las preguntas de investigación, la metodología, los casos ilustrativos, el análisis interpretativo y la discusión, seguidos de limitaciones y recomendaciones operativas para agencias, traductores y clientes en un ecosistema multilingüe marcado por la IA.

2. Marco teórico y antecedentes

La Gestión de Proyectos de Traducción (GPT) ha experimentado una evolución significativa impulsada por cambios tecnológicos, transformaciones operativas y el crecimiento exponencial de los flujos multilingües. En sus inicios, la GPT se concebía como una función administrativa centrada en la asignación de tareas, la coordinación de proveedores y la supervisión básica de entregas. Sin embargo, el aumento del volumen de contenido digital, la necesidad de escalabilidad y la diversificación de mercados globales reconfiguraron este rol hacia un perfil altamente especializado que integra conocimientos lingüísticos, tecnológicos y de gestión (Antúnez Brizuela, 2019; Fuentes-Pérez & Cano Fernández, 2025). En esta línea, trabajos recientes resaltan que el gestor de proyectos no solo coordina recursos, sino que diseña y optimiza flujos de trabajo, actúa como mediador entre actores y tecnologías, y administra riesgos derivados de la complejidad organizacional y discursiva de los encargos multilingües (Rogl & Risku, 2024). De este modo, el gestor pasó a desempeñar funciones de ingeniería del flujo de trabajo, gestión del riesgo, control de calidad multicapa — entendido como un aseguramiento de la calidad en varios niveles complementarios (por ejemplo, controles automatizados con herramientas de QA (*Quality*

Assurance / aseguramiento de la calidad), revisión lingüística y terminológica humana, verificación de guías de estilo del cliente y validación funcional del entregable en su formato final) — y comunicación estratégica con clientes y equipos de traducción dispersos, es decir, equipos geográficamente distribuidos (a menudo remotos) que operan en husos horarios distintos y requieren coordinación, trazabilidad y alineación de criterios.

El desarrollo de sistemas de traducción automática neuronal (NMT), los modelos de lenguaje a gran escala (LLM) y las plataformas colaborativas ha introducido un cambio estructural en los flujos de trabajo. Este fenómeno, frecuentemente denominado *AI-shift*, ha modificado las expectativas de clientes y proveedores, intensificando la tensión entre costo, velocidad y calidad (Das, 2018; Mohamed et al., 2024). En este contexto, los modelos híbridos —que combinan traducción humana, traducción automática (TA) y/o inteligencia artificial (IA), posedición y control de calidad automatizado— se han consolidado como prácticas predominantes en sectores como *e-commerce*, tecnología y localización (Hiebl & Gromann, 2023). Diversos estudios señalan que esta convergencia humano-máquina exige nuevas competencias profesionales, tales como alfabetización tecnológica, capacidad crítica para evaluar resultados automáticos, comprensión operativa de estándares y habilidades de comunicación para gestionar expectativas de clientes con distintos niveles de madurez tecnológica (Cadwell et al., 2016; Krüger, 2025). Desde esta perspectiva, el gestor de proyectos opera como un arquitecto del flujo lingüístico, responsable de diseñar procesos eficientes, prever riesgos operativos y garantizar coherencia cultural y terminológica a lo largo del proyecto.

En paralelo, investigaciones recientes han reforzado el lugar de la GPT como campo de análisis en sí mismo, al destacar la interacción entre prácticas de traducción y enfoques de gestión de proyectos en entornos multilingües complejos. Al-Tarawneh (2025), Tarawneh (2025) y Al-Tarawneh y Al-Badawi (2025) muestran que la eficiencia en proyectos multilingües no depende únicamente del desempeño lingüístico individual, sino de decisiones de planificación, asignación, seguimiento y control que influyen directamente en la calidad final y en el cumplimiento de plazos. En consecuencia, un enfoque que articule dimensiones lingüísticas y operativas permite comprender la

traducción profesional como una actividad situada, condicionada por variables organizacionales, tecnológicas y económicas.

La adopción masiva de la IA también ha replanteado los fundamentos de la gestión terminológica, es decir, el conjunto de prácticas destinadas a identificar, consensuar, registrar, mantener y aplicar de manera coherente los términos especializados de un dominio (por ejemplo, mediante glosarios, bases terminológicas, memorias de traducción y guías de estilo), históricamente considerada un componente estable del proceso. La literatura reciente advierte que los sistemas de NMT y, en determinados escenarios, los LLM aplicados a traducción pueden favorecer inconsistencias y generar *drift* terminológico, entendido como la desviación gradual respecto de la terminología validada para un proyecto, especialmente cuando no se implementan restricciones explícitas o mecanismos de control robustos (Massion, 2021; Moslem et al., 2023; Kim et al., 2024). Asimismo, investigaciones sobre traducción automática y posedición han señalado que los sistemas actuales pueden producir resultados lingüísticamente fluidos, pero no necesariamente consistentes en términos de precisión conceptual y estabilidad terminológica, lo cual vuelve indispensable mantener intervenciones humanas estratégicas (Koponen, 2016; Bowker & Buitrago Ciro, 2019). En una línea convergente, los estudios sobre alucinaciones en contextos automáticos han documentado salidas plausibles, pero semánticamente inexactas, que pueden introducir equivalentes técnicamente inadecuados o unidades léxicas no respaldadas por el texto fuente (Guerreiro et al., 2023; Xu et al., 2023). Estas limitaciones, asociadas al funcionamiento probabilístico de la IA, convierten la gestión terminológica en un mecanismo crítico de control de calidad en dominios sensibles como el jurídico, el médico-regulatorio o el técnico especializado (Moneus & Sahari, 2024).

En consecuencia, agencias y equipos profesionales suelen recurrir a flujos híbridos que incluyen extracción automática de términos, consolidación de candidatos, validación experta, glosarios dinámicos y controles de congruencia integrados en herramientas CAT (Computer-Assisted Translation) y en sistemas de QA. Este enfoque combina velocidad con confiabilidad, pero mantiene la supervisión humana como componente ético y operativo indispensable. En este marco, Abed (2025) subraya que la transformación digital del sector

también implica desafíos metodológicos, tales como la trazabilidad de decisiones, la gobernanza terminológica, entendida como el conjunto de políticas, roles, procedimientos y herramientas destinadas a definir, validar, actualizar y hacer cumplir de manera consistente la terminología aprobada (por ejemplo, mediante glosarios, bases terminológicas y controles de QA) a lo largo de todo el flujo de trabajo, y la consistencia en proyectos escalables, mientras que revisiones recientes coinciden en que la integración de IA puede incrementar la productividad, aunque persisten limitaciones vinculadas con sesgos, precisión y control de calidad en tareas especializadas (Mohamed et al., 2024; Shahmerdanova, 2025). Como apoyo ilustrativo, se adjuntan en los Anexos A–C ejemplos de glosarios gobernados y reglas de QA por criticidad, utilizados como referencia conceptual para describir el control terminológico en flujos híbridos.

Finalmente, la calidad en entornos híbridos debe entenderse como un constructo multidimensional, que integra dimensiones lingüísticas, funcionales, culturales y operativas, junto con requisitos específicos del cliente. Los estándares ISO 17100 e ISO 18587, las taxonomías de error del marco MQM (en inglés, *Multidimensional Quality Metrics*), un sistema estandarizado para evaluar la calidad mediante categorías jerárquicas de error (p. ej., *Accuracy, Fluency, Terminology, Style, Locale*, etc.) y las herramientas automatizadas de control de calidad permiten construir sistemas robustos que reducen la propagación de errores y optimizan la toma de decisiones (García & Pena, 2021). Sin embargo, como advierten Floridi y Chiriatti (2020), el comportamiento opaco y en ocasiones impredecible de los LLM refuerza la necesidad de supervisión humana estratégica, especialmente en la evaluación de hechos, matices culturales y terminología crítica. En esta línea, se ha sostenido que la evaluación debe ampliarse más allá de unidades aisladas, atendiendo a coherencia discursiva y consistencia conceptual a nivel textual y documental (Läubli et al., s. f.). En el caso de la traducción audiovisual y la localización multimodal, además, se añaden restricciones temporales, pragmáticas y semióticas que limitan la automatización total (Georgakopoulou, 2019; Davitti, 2025). Así, el gestor de proyectos se sitúa en el centro de un modelo de calidad continuo y cíclico que integra IA, conocimiento experto y retroalimentación permanente, en

un campo profesional marcado por tensiones entre eficiencia productiva y sostenibilidad laboral (Moorkens, 2017; TAUS, 2023).

3. Preguntas de investigación

A partir del marco teórico y del propósito del estudio, se plantean las siguientes preguntas de investigación:

1. ¿Qué dimensiones (lingüísticas, terminológicas, funcionales y operativas) deben integrarse para evaluar flujos de trabajo híbridos?
2. ¿Cómo impacta el *AI-shift* en el equilibrio calidad–tiempo–costo en proyectos multilingües?
3. ¿Qué rol cumplen la gestión terminológica y los controles multicapa en la mitigación de riesgos (*drift*, inconsistencias, alucinaciones) en estos flujos?

4. Metodología

Este estudio adopta un enfoque metodológico híbrido orientado al desarrollo de un modelo conceptual aplicable a contextos reales de la Gestión de Proyectos de Traducción. La metodología combina: (a) una revisión documental sistemática, (b) observación indirecta de la práctica profesional y análisis comparativo, y (c) la construcción de estudios de caso ilustrativos basados en patrones recurrentes observables en agencias de traducción pequeñas y medianas. La construcción de los casos se sustenta en la convergencia entre los hallazgos de la revisión documental y los patrones operativos identificados a partir de la observación indirecta. La elección de este enfoque responde a la necesidad de integrar evidencia conceptual, regularidades de la práctica y escenarios plausibles en un campo altamente dinámico en el que la confidencialidad y la heterogeneidad de los flujos suelen limitar el acceso a datos comparables en entornos reales. Con el fin de operacionalizar los insumos analizados, se elaboró una matriz de categorías que permitió estructurar y comparar los componentes centrales del flujo de trabajo (por ejemplo, etapas del proceso, roles involucrados, puntos de control de calidad, riesgos frecuentes y mecanismos de mitigación). Esta matriz constituyó la base para diseñar los casos ilustrativos y para fundamentar el análisis del modelo propuesto.

4.1. Participantes y contexto del estudio

La investigación no se basa en la intervención directa de individuos, sino en la observación indirecta de agencias activas en los mercados europeo, asiático y latinoamericano que operan con flujos tradicionales, híbridos y centrados en la IA. Los actores involucrados incluyen gestores de proyectos, traductores humanos, poseditores, terminólogos, revisores senior y especialistas en control de calidad automatizado. Esta diversidad permitió identificar patrones de decisión y estrategias de mitigación aplicadas de manera integrada en entornos de producción.

4.2. Instrumentos y corpus de análisis

Se utilizaron tres fuentes de información complementaria. Primero, se llevó a cabo una revisión bibliográfica (2020–2023) centrada en traducción automática (NMT), modelos de lenguaje a gran escala (LLMs), posedición, gestión terminológica y estándares de calidad como ISO 17100, ISO 18587 y MQM. Segundo, se analizaron documentos operativos de agencias, que incluyeron glosarios, configuraciones de motores de traducción automática, inteligencia artificial, reportes de control de calidad y registros de *drift* terminológico. Tercero, se trabajó con un corpus compuesto por instrucciones del cliente, segmentación en las herramientas de traducción, informes de incongruencias y ejemplos reales de divergencias terminológicas, con el objetivo de contextualizar los riesgos y decisiones operativas observadas en escenarios operativos. Como apoyo ilustrativo, se adjuntan en los Anexos A–C ejemplos de glosarios gobernados y reglas de QA por criticidad, utilizados como referencia conceptual para describir el control terminológico en flujos híbridos.

4.2.1. Cuadro analítico de categorías para el análisis comparativo

Con el propósito de sistematizar la información obtenida a partir de la revisión documental y de la observación indirecta de la práctica profesional, se elaboró la Tabla 1 que presenta un cuadro analítico de categorías que permitió organizar los datos en dimensiones comparables. Este instrumento funcionó como una estructura de codificación y síntesis, orientada a identificar regularidades operativas en flujos tradicionales, híbridos y centrados en la IA,

comparar patrones de decisión y control de calidad, y fundamentar el diseño del modelo multidimensional y de los estudios de caso ilustrativos presentados en las secciones posteriores.

Categoría analítica	**Descripción**	**Indicadores observables en la práctica/ documentos**	**Función en el estudio**
Arquitectura del flujo	Organización del proceso en etapas y secuencias de trabajo	Recepción del encargo, análisis, preproducción, producción, QA, entrega, retroalimentación	Permite comparar modelos de flujo (Tabla 2) y describir dónde se concentran las decisiones críticas
Roles y responsabilidades	Distribución de tareas entre actores del proceso	PM, traductor, poseditor, terminólogo, revisor senior, especialista QA	Visibiliza redistribución de funciones y dependencia de perfiles expertos en entornos con IA
Puntos de control de calidad	Momentos y mecanismos de control a lo largo del flujo	QA preventivo, control en proceso, QA final; revisión bilingüe/ monolingüe; validación funcional	Sustenta el concepto de calidad multicapa y la supervisión humana estratégica
Riesgos recurrentes	Problemas frecuentes asociados a calidad y confiabilidad	*Drift* terminológico, inconsistencias, alucinaciones, errores factuales, pérdida de estilo/voz	Justifica la inclusión de la dimensión "riesgo" en el modelo multidimensional (Tabla 3)
Mecanismos de mitigación	Acciones y recursos	Glosarios, guías de estilo,	Orienta el diseño de casos

	para reducir riesgos	restricciones terminológicas, MQM/ISO, CAT tools y QA tools, revisión especializada	plausibles y permite describir estrategias de control en cada flujo

Tabla 1: Cuadro analítico de categorías empleado en el estudio.

Este cuadro analítico permitió construir una base de comparación consistente entre los tres modelos de flujo de trabajo y operacionalizar las dimensiones que conforman el modelo multidimensional propuesto. A partir de esta sistematización, se desarrolló el análisis comparativo presentado en la sección 4.3 y se diseñaron los casos ilustrativos que permiten ejemplificar decisiones típicas de gestión en contextos mediados por IA.

4.3 Diseño del estudio

El diseño metodológico integra tres componentes analíticos. El primero consistió en un análisis comparativo estructurado de los tres modelos de flujo de trabajo predominantes en la industria: tradicional, híbrido y centrado en la IA. Este análisis consideró características operativas, fortalezas, limitaciones y dominios de aplicabilidad, con el fin de evaluar la idoneidad en términos de calidad, costo y escalabilidad. La Tabla 1 sintetiza esta comparación.

Modelo	**Características clave**	**Fortalezas**	**Limitaciones**	**Mejor aplicabilidad**
Tradicional	Traducción humana; revisión bilingüe y monolingüe; control estilístico profundo	Máxima precisión cultural; estilo cuidado; ausencia de *drift* terminológico	Costo elevado; tiempos extensos; baja escalabilidad	Jurídico, literario, médico-regulatorio
Híbrido	TA/IA como base; posedición; control de calidad	Alta productividad; equilibrio costo–calidad;	Requiere poseditores expertos; variabilidad en	*E-commerce*, tecnología, localización

	automatizado e intervención humana	reducción de tiempos	textos complejos	
Centrado en la IA	LLM como núcleo; uso de *prompts*, guías de estilo y recursos terminológicos; control predictivo	Velocidad y escalabilidad; adaptación estilística en tareas de bajo riesgo	Alucinaciones; incongruencias; riesgos éticos y terminológicos	Alto volumen, borradores, contenido dinámico no sensible o no publicable

Tabla 2: Comparación de modelos de flujo de trabajo en la gestión de proyectos de traducción.

A partir del análisis comparativo, sustentado en el cuadro analítico de categorías, se desarrolló un modelo multidimensional destinado a explicar las variables que determinan la efectividad operativa y los riesgos asociados a cada modelo. Este modelo integra dimensiones de calidad, congruencia terminológica, escalabilidad, costo, intervención humana, complejidad temática y adecuación a los requisitos del cliente. La Tabla 3 presenta las dimensiones clave del modelo y su relevancia práctica.

Dimensión	**Descripción**	**Indicadores observados**	**Impacto de la IA / Riesgos**	**Relevancia operativa**
Calidad lingüística	Fidelidad semántica y estilística	Tipología de errores (MQM); tasa de error	Alucinaciones; errores factuales	Evaluación del producto final
Congruencia terminológica	Uso de términos validados	*Drift*; sinónimos no aprobados y no deseables	Terminología inventada; variabilidad	Gestión del riesgo en dominios críticos
Escalabilidad	Capacidad de procesar volumen	Velocidad	Reducción del control por el uso de la IA	Localización masiva, ideal para sitios de *e-commerce*
Costo / Esfuerzo	Tarifas de posedición	PEE; horas invertidas	Presión tarifaria; precarización	Sostenibilidad del proveedor

Intervención humana	Nivel de supervisión requerido	CC preventivo, durante el proceso y final	Intervención humana obligatoria	Mitigación de errores críticos
Complejidad del dominio	Sensibilidad temática	Requisitos regulatorios y técnicos	IA no confiable en sectores críticos	Medicina, jurídico, técnico
Adecuación al cliente	Estilo, voz de marca y registro	Cumplimiento de guías	Homogeneización / pérdida de clientes	Branding y contenido creativo

Tabla 3: Modelo multidimensional para analizar flujos de trabajo en gestión de proyectos de traducción.

A modo complementario y con fines ilustrativos, la figura 1 esquematiza el Modelo Multidimensional GPT y sintetiza visualmente las relaciones jerárquicas entre las dimensiones analizadas en la Tabla 3. El diagrama organiza estas dimensiones en distintos niveles para mostrar cómo interactúan en un flujo de trabajo híbrido mediado por la IA.

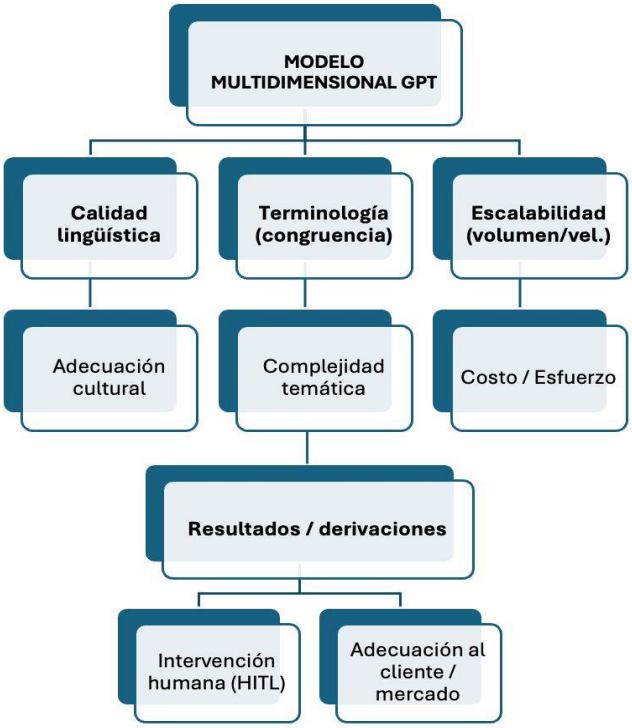


Figura 1: Modelo multidimensional para la gestión de proyectos de traducción.

Este esquema permite visualizar cómo las dimensiones superiores (calidad lingüística, congruencia terminológica y escalabilidad) funcionan como indicadores de la robustez operacional del proceso, mientras que las dimensiones intermedias (adecuación cultural, complejidad temática y costo/esfuerzo) operan como filtros críticos que condicionan las decisiones operativas. Finalmente, las dimensiones inferiores (intervención humana y adecuación al cliente/mercado) determinan la viabilidad real del flujo en contextos profesionales, ya que establecen hasta qué punto el proceso puede implementarse sin comprometer estándares de calidad, seguridad o coherencia con los requisitos del cliente.

4.4 Procedimientos

El procedimiento se desarrolló en tres etapas. En primer lugar, se sistematizó la revisión documental con el fin de identificar y organizar las categorías analíticas que sustentan el modelo multidimensional. Estas categorías se operacionalizaron mediante un cuadro analítico compuesto por cinco ejes comparables: (i) arquitectura del *workflow* (etapas y secuencias), (ii) roles y responsabilidades, (iii) puntos de control de calidad, (iv) riesgos recurrentes (p. ej., *drift* terminológico, inconsistencias y alucinaciones) y (v) mecanismos de mitigación (p. ej., glosarios, guías de estilo, restricciones terminológicas y QA automatizado).

En segundo lugar, a través de observación indirecta de la práctica profesional, se identificaron estrategias recurrentes de configuración de flujos y de uso de herramientas de traducción y control de calidad, tales como: selección del modelo de producción (tradicional/híbrido/centrado en IA) según el dominio y el nivel de riesgo; integración de recursos terminológicos (glosarios y bases) en entornos CAT; aplicación de controles automatizados de QA; escalamiento de revisión humana especializada en dominios críticos; y validación final orientada a requisitos del cliente y adecuación funcional del entregable.

En tercer lugar, se construyeron cuatro estudios de caso ilustrativos inspirados en patrones observables en agencias pequeñas y medianas, con el propósito de aplicar el modelo y examinar su utilidad explicativa en escenarios jurídicos, técnicos, audiovisuales y de e-commerce. Los casos se analizaron cualitativamente articulando el corpus observado, el marco teórico y las dimensiones del modelo, con el objetivo de mostrar cómo se distribuyen decisiones, riesgos y controles a lo largo del flujo de trabajo.

5. Estudios de caso

Con el fin de operacionalizar el modelo multidimensional y examinar su utilidad interpretativa en escenarios plausibles, se desarrollaron cuatro estudios de caso simulados basados en patrones documentados en la práctica profesional de agencias multilingües. En estudios de traducción y en análisis de flujos de trabajo de la industria lingüística, el uso de diseños cualitativos y estudios de caso constituye un recurso metodológico pertinente para representar decisiones

típicas, explorar tensiones operativas y describir configuraciones recurrentes, particularmente cuando el acceso a datos comparables se ve restringido por condiciones propias del entorno profesional (por ejemplo, confidencialidad, restricciones comerciales y variabilidad de proyectos) (O'Brien & Mellinger, 2025; Risku et al., 2019; Saldanha & Brien, 2013). En esta línea, investigaciones socio-cognitivas sobre traducción en el lugar de trabajo han mostrado el valor de observar procesos y redes de interacción en contextos reales para comprender cómo evolucionan los roles, los flujos y la toma de decisiones en entornos mediados por tecnología (Risku & Windhager, 2013; Risku, Rogl & Milošević, 2019).

La construcción de estos casos sigue una lógica de representatividad estructural: cada uno ejemplifica un dominio donde el *AI-shift* introduce desafíos específicos que requieren decisiones estratégicas por parte del gestor de proyectos. Los casos no pretenden reproducir datos empíricos ni describir situaciones reales, sino aislar dinámicas características y mostrar cómo las dimensiones del modelo —calidad, congruencia terminológica, escalabilidad, costo/esfuerzo, intervención humana, complejidad temática y adecuación cultural— interactúan en contextos operativos concretos. Cada caso incorpora un conflicto principal, una respuesta operativa y un *insight* que conecta el escenario con el modelo.

Caso 1: Proyecto jurídico y desviaciones terminológicas en flujo híbrido

Una agencia recibe un proyecto jurídico de 90.000 palabras con entrega urgente. Para acelerar el proceso, se utiliza un flujo híbrido con pretraducción mediante NMT/IA, pero la primera revisión evidencia incongruencias terminológicas y variación en equivalentes críticos, atribuibles a restricciones terminológicas insuficientes y a una configuración subóptima del motor (por ejemplo, falta de integración efectiva de glosarios o parámetros de consistencia). Dado que la sensibilidad del dominio exige precisión terminológica estricta, se implementa un macroglosario validado por un lingüista experto, acompañado de controles automatizados de QA en dos capas y alineación terminológica sistemática. Este caso muestra que, en dominios regulados, la congruencia terminológica y la intervención humana especializada constituyen dimensiones críticas del modelo

multidimensional, y que incluso un flujo híbrido puede requerir reforzar mecanismos tradicionales para garantizar cumplimiento normativo.

Caso 2: *E-commerce* y optimización mediante el flujo híbrido

Un cliente solicita la traducción semanal de 40.000 a 60.000 palabras en ocho idiomas, con ciclos de entrega particularmente breves. Para equilibrar costo, velocidad y consistencia estilística, la agencia implementa un flujo híbrido basado en TA/IA seguida de posedición y control de calidad integrado en la herramienta CAT. En términos operativos, este tipo de configuración suele asociarse con incrementos relevantes de productividad y reducciones de costo por palabra, en la medida en que se preserve una posedición experta y una gobernanza terminológica estable. Este caso ejemplifica la dimensión de escalabilidad del modelo multidimensional, mostrando que los flujos híbridos pueden sostener grandes volúmenes con control de calidad aceptable siempre que existan puntos de verificación consistentes y posedición especializada.

Caso 3: Subtitulación audiovisual y límites de la IA generativa

Una plataforma audiovisual experimenta con subtítulos generados mediante IA. Aunque el sistema resuelve adecuadamente equivalencias básicas, se observan dificultades en sincronía**,** entendida como la adecuación temporal entre el texto subtitulado y el habla en pantalla (p. ej., segmentación, duración del subtítulo, velocidad de lectura y alineación con entradas y salidas del audio). Asimismo, se detectan problemas en matices humorísticos, tonos pragmáticos y referencias culturales. En consecuencia, la intervención humana se vuelve indispensable para restaurar claridad, naturalidad y adecuación audiovisual, apoyada en guías de estilo y criterios de segmentación. Este caso ilustra cómo la dimensión de adecuación cultural y funcional del modelo no puede delegarse plenamente a la IA en productos multimodales, dada su tendencia a homogeneizar el estilo y producir aproximaciones semánticas que no siempre son compatibles con la pragmática del discurso audiovisual.

Caso 4: Proyecto médico-regulatorio y retorno forzado al modelo tradicional

Un fabricante de dispositivos médicos requiere traducciones urgentes en todos los idiomas de la Unión Europea. Aunque la IA ofrece buenos resultados iniciales en algunos idiomas, el contenido técnico altamente regulado evidencia un riesgo significativo de errores conceptuales. Los resultados varían de manera incongruente entre lenguas, y en ciertos segmentos, el sistema genera términos plausibles pero incorrectos. Ante el riesgo legal, la agencia abandona el flujo híbrido y retorna al modelo tradicional, priorizando la intervención humana experta y cumpliendo con las normativas europeas que requieren traducción y revisión. Este caso evidencia que, en dominios críticos, la dimensión de complejidad temática determina el nivel de intervención humana necesaria, reafirmando la imposibilidad de adoptar flujos centrados en la IA sin una capa de control especializada.

Estos cuatro casos constituyen material ilustrativo que permite visualizar cómo el modelo multidimensional articula calidad, terminología precisa, escalabilidad, costo/esfuerzo y supervisión humana en distintos contextos. Asimismo, ofrecen un puente conceptual entre la metodología y el análisis posterior, donde se discuten implicancias operativas del *AI-shift* para la gestión de proyectos de traducción.

6. Resultados analíticos y recomendaciones derivadas de la aplicación del modelo multidimensional

El presente apartado reporta los resultados analíticos obtenidos a partir de la aplicación del modelo multidimensional a los cuatro estudios de caso ilustrativos. Dado el carácter conceptual y representacional del diseño, estos resultados no se interpretan como evidencia empírica estadísticamente verificable, sino como hallazgos interpretativos sobre patrones de decisión, tensiones operativas y relaciones plausibles entre dimensiones (calidad, congruencia terminológica, escalabilidad, costo/esfuerzo, intervención humana, complejidad del dominio y adecuación al cliente). En este marco, los resultados se expresan como regularidades consistentes con la literatura reciente sobre IA en traducción y posedición (Guerreiro et al., 2023; Moslem et al., 2023; Strandvik, 2025).

6.1 Resultados 1: Efectos del modelo de flujo de trabajo sobre calidad, terminología y escalabilidad

El análisis comparativo evidencia que el tipo de flujo de trabajo actúa como variable moduladora de las dimensiones críticas del modelo. En primer lugar, los flujos tradicionales tienden a maximizar el control lingüístico y la estabilidad terminológica, especialmente en dominios regulados, donde la revisión experta y los procedimientos de aseguramiento de calidad resultan determinantes (Strandvik, 2025). En segundo lugar, los flujos híbridos muestran mayor potencial para sostener volumen y plazos ajustados sin comprometer necesariamente la consistencia, siempre que se mantenga posedición experta y control terminológico integrado (ISO, 2017; Moslem et al., 2023). En tercer lugar, los flujos centrados en IA, aun cuando ofrecen ventajas de velocidad y escalabilidad, incrementan la variabilidad en precisión, estilo y terminología, lo que exige gobernanza robusta y supervisión humana estratégica (Guerreiro et al., 2023).

En términos interpretativos, la aplicación del modelo permite identificar un principio operativo consistente: la selección del flujo no debe basarse en la tecnología disponible, sino en la combinación de dimensiones relevantes del encargo. En proyectos de alta sensibilidad conceptual o regulatoria, las dimensiones de congruencia terminológica, complejidad del dominio e intervención humana emergen como prioritarias frente a la velocidad. En cambio, en contextos de alto volumen, la escalabilidad y el costo/esfuerzo adquieren centralidad, sin eliminar la necesidad de control de calidad multicapa y gestión terminológica activa (Strandvik, 2025).

Los gestores podrían adoptar un criterio de "ajuste por dominio y riesgo", seleccionando configuraciones de flujo en función de sensibilidad conceptual, volumen, requisitos de estilo, riesgo regulatorio y expectativas del cliente.

6.2. Resultado 2: Riesgos potenciales del AI-shift y eficacia de los mecanismos de mitigación

Los casos analizados permiten identificar riesgos recurrentes asociados a la integración de IA: *drift* terminológico, inconsistencias, errores factuales y salidas plausibles, pero semánticamente inadecuadas, además de pérdida de matices pragmáticos y homogeneización estilística. Estos riesgos han sido documentados en

la investigación reciente sobre traducción automática y modelos multilingües, particularmente en condiciones de dominio especializado o cuando se requiere control estricto de adecuación semántica (Guerreiro et al., 2023). En términos terminológicos, la literatura sobre restricción terminológica y control de terminología en MT/LLMs muestra que el cumplimiento consistente de glosarios no es automático y requiere mecanismos específicos de integración o decodificación restringida (Dinu et al., 2019; Hasler et al., 2018; Moslem et al., 2023).

Desde el modelo multidimensional, tres dimensiones operan como filtros críticos para mitigar riesgos: (i) intervención humana estratégica, (ii) congruencia terminológica y (iii) adecuación cultural y funcional. El análisis sugiere que la combinación de glosarios validados, guías de estilo, controles automáticos de QA y revisión experta incrementa la estabilidad del producto y reduce la propagación de errores en proyectos con IA, coherente con enfoques de calidad en capas y con estándares de posedición (ISO, 2017; Strandvik, 2025).

Toda implementación de IA debería operar bajo un enfoque *human-in-the-loop*, integrando validación terminológica en capas, guías de estilo explícitas y control de calidad continuo, especialmente en dominios sensibles (ISO, 2017).

6.3. Resultado 3: Rol estratégico del gestor de proyectos y sostenibilidad operativa del flujo

Los casos ilustrativos confirman que el gestor de proyectos cumple una función estructural en la mitigación de riesgo y en la eficiencia del flujo. La selección del modelo operativo, la parametrización del proceso (p. ej., integración terminológica), el diseño de puntos de control y la definición del nivel de posedición determinan el equilibrio entre calidad, escalabilidad y costo/esfuerzo. En este sentido, el modelo multidimensional funciona como guía para la toma de decisiones, al permitir priorizar dimensiones según el dominio y anticipar dónde se concentran los riesgos.

Asimismo, el análisis del componente costo/esfuerzo sugiere que una integración acrítica de IA puede incrementar la carga cognitiva de la posedición y tensionar la sostenibilidad del proceso, particularmente cuando se subestima el esfuerzo real requerido para

detectar errores, asegurar consistencia y restaurar estilo (Koponen, 2012; Rico Pérez, 2024; Vieira, 2014). La evidencia disponible sobre esfuerzo cognitivo en posedición y traducción colaborativa humano-máquina respalda que la productividad no siempre se traduce linealmente en menor carga mental o menor esfuerzo de revisión, lo que afecta tanto calidad como condiciones profesionales (Chen, 2025).

El gestor debería actuar como arquitecto del flujo, en el que articularía tecnología, control experto y requisitos del cliente, e incluiría criterios explícitos para tarifación de posedición acorde al esfuerzo real y restricciones claras sobre automatización en dominios críticos (ISO, 2017; Rico Pérez, 2024).

En síntesis, la aplicación del modelo multidimensional muestra que la efectividad de los flujos de trabajo no depende exclusivamente de la tecnología utilizada, sino del equilibrio entre calidad, consistencia terminológica, escalabilidad, costo/esfuerzo, supervisión humana, complejidad temática y adecuación a los requisitos del cliente. Los casos ilustrativos permiten observar que los flujos híbridos, cuando se implementan con gobernanza terminológica y QA en capas, constituyen una opción operativamente robusta en entornos reales, al combinar eficiencia con mecanismos de control ajustables al nivel de riesgo del contenido (ISO, 2017; Strandvik, 2025).

7. Discusión

Esta discusión articula el marco teórico, el modelo multidimensional y los cuatro casos ilustrativos para sintetizar qué condiciones hacen más adecuado cada flujo de trabajo (tradicional, híbrido o centrado en IA) y qué dimensiones deben priorizarse en la toma de decisiones. En los escenarios regulados (jurídico y médico-regulatorio), la evidencia conceptual converge en que el modelo tradicional ofrece el mayor control sobre precisión conceptual y estabilidad terminológica, lo cual es consistente con la literatura sobre variabilidad terminológica y errores en dominios especializados cuando se reduce la supervisión humana (Bowker & Buitrago Ciro, 2019; Koponen, 2016). En cambio, en contextos de alto volumen (e-commerce), el flujo híbrido aparece como la opción más robusta, ya que permite sostener escalabilidad y plazos exigentes sin degradar necesariamente la calidad si se mantiene posedición experta y control

sistemático (Pinnis et al., 2016; Sánchez-Gijón et al., 2019; Hiebl & Gromann, 2023). Por contraste, el uso centrado exclusivamente en IA maximiza velocidad, pero tiende a incrementar la variabilidad en precisión factual y consistencia terminológica, especialmente en contenidos sensibles, lo que hace inviable su adopción sin capas de control (Floridi & Chiriatti, 2020; Qin et al., 2024). Esta comparación no es lineal ni "tecnología vs. tradición": los casos muestran decisiones aparentemente contradictorias, pero operativamente coherentes (p. ej., proyectos con IA que requieren retroceder hacia controles tradicionales), lo que refuerza la idea de que la idoneidad depende del riesgo del dominio y de los requisitos del cliente, no del recurso tecnológico disponible (Antúnez Brizuela, 2019; Fuentes-Pérez & Cano Fernández, 2025).

La aplicación del modelo permite, además, delimitar qué variables explican mejor el desempeño esperado de cada flujo y por qué ciertos problemas se intensifican con el AI-shift. Los casos sostienen que la congruencia terminológica, la intervención humana estratégica y la complejidad temática operan como ejes de control cuando se trabaja con NMT/LLMs, dado que el drift terminológico, las inconsistencias y las salidas plausibles pero inexactas han sido reportadas como riesgos reales en sistemas automáticos, en particular si no existen restricciones terminológicas explícitas o mecanismos robustos de verificación (Massion, 2021; Moslem et al., 2023; Kim et al., 2024; Guerreiro et al., 2023; Xu et al., 2023). En términos operativos, esto se traduce en que la "calidad" no puede tratarse como un atributo único, sino como un constructo multidimensional que requiere aseguramiento en capas (p. ej., QA automatizado + revisión humana + validación terminológica), tal como sugieren marcos normativos y de evaluación de errores (ISO 17100; ISO 18587; MQM; García & Pena, 2021). El caso audiovisual refuerza un límite complementario: aun con outputs fluidos, la IA presenta restricciones cuando la adecuación depende de sincronía, humor, pragmática y referencias culturales, en línea con los límites descritos para automatización en subtitulado (Georgakopoulou, 2019). Así, el aporte central del modelo no es "probar" superioridad de una modalidad, sino ofrecer un mapa de priorización de dimensiones que permite anticipar trade-offs realistas entre escalabilidad, costo/esfuerzo y control experto.

Finalmente, los cuatro casos confirman que el gestor de proyectos es un componente estructural del sistema, ya que su toma de decisiones regula el equilibrio entre eficiencia y riesgo. La literatura ya caracteriza al gestor como diseñador del flujo, mediador entre tecnología y estándares, y responsable de coordinar control de calidad, terminología y expectativas del cliente (Antúnez Brizuela, 2019; Cadwell et al., 2016; Rogl & Risku, 2024). En este marco, los hallazgos analíticos sugieren que la sostenibilidad del proceso depende de políticas explícitas de posedición y revisión (O'Brien, 2021, 2022) y de evitar la automatización total en dominios críticos, donde la supervisión humana sigue siendo indispensable (Läubli et al., 2020). Asimismo, las tensiones asociadas a presión tarifaria y sobrecarga en posedición son coherentes con reportes y estudios sobre precarización cuando se integra TA/IA sin reconocer el esfuerzo real del trabajo (Moorkens, 2017; TAUS, 2023), lo que refuerza la necesidad de criterios de asignación, tarifación y control acordes al riesgo del contenido. En conjunto, el modelo multidimensional permite explicar por qué el flujo híbrido emerge como configuración "óptima" en escenarios frecuentes, sin excluir que, bajo alta sensibilidad conceptual, el retorno a mecanismos tradicionales sea la decisión más responsable: la eficacia no está en la IA por sí misma, sino en cómo se articula con control experto, gobernanza terminológica y QA multicapa.

8. Limitaciones, proyecciones futuras e implicancias profesionales

Este estudio presenta limitaciones propias de un diseño conceptual basado en casos ilustrativos y observación indirecta, por lo que sus aportes deben interpretarse como hallazgos analíticos plausibles más que como evidencia empírica cuantificable: no permite inferencias estadísticas ni medir de forma objetiva la variabilidad asociada a pares lingüísticos, configuraciones específicas de NMT/LLMs, carga cognitiva de posedición o magnitud real del *drift* terminológico. En consecuencia, se requiere investigación futura con datos comparables (corpus, métricas de calidad, registros de QA, análisis longitudinales y observación en entornos productivos) para evaluar empíricamente la interacción entre IA, supervisión humana y terminología crítica. Aun así, el modelo multidimensional ofrece implicancias profesionales claras: confirma que la supervisión humana sigue

siendo indispensable para garantizar precisión conceptual, coherencia terminológica y adecuación cultural, y que la competencia tecnológica y el conocimiento del dominio se vuelven habilidades estratégicas para traductores y poseditores. Para los gestores de proyectos, el modelo refuerza un rol de diseño y gobernanza del flujo (esto es, la definición de reglas, responsabilidades, puntos de control y protocolos de decisión que organizan cómo se traduce, se revisa y se valida un entregable en cada etapa) —selección de modalidad según riesgo y requisitos, QA multicapa, tarifación acorde al esfuerzo real y orientación responsable a clientes—, mientras que para agencias y organizaciones destaca que la sostenibilidad depende de evitar automatización acrítica e implementar capacitación, gobernanza terminológica y control de calidad robusto en entornos híbridos.

9. Conclusiones

Este capítulo aporta al campo de la Gestión de proyectos de traducción en la era de la IA mediante la formulación de un modelo multidimensional que integra variables lingüísticas, operativas y tecnológicas para evaluar la idoneidad de flujos de trabajo tradicionales, híbridos y centrados en IA. El modelo articula siete dimensiones críticas —calidad lingüística, congruencia terminológica, escalabilidad, costo/esfuerzo, intervención humana, complejidad temática y adecuación al cliente— y permite comprender que la eficacia del proceso no depende exclusivamente de la tecnología empleada, sino del equilibrio estratégico entre dichas dimensiones según el dominio y el riesgo del encargo. A partir de casos ilustrativos basados en patrones profesionales, el análisis muestra que los flujos híbridos resultan operativamente robustos en escenarios de alto volumen, mientras que en dominios regulados la supervisión humana experta y la gobernanza terminológica siguen siendo condiciones indispensables para mitigar riesgos como *drift* terminológico, inconsistencias y salidas plausibles pero inexactas. En conjunto, el modelo ofrece una herramienta analítica útil para orientar decisiones de diseño y gobernanza del flujo, y establece una base conceptual para futuras investigaciones empíricas sobre calidad, terminología y gestión del riesgo en entornos mediados por IA.

Referencias

Abed, L. (2025). The impact of AI applications on translation and terminology management: Digital transformation and methodological challenges. *Cahiers de Traduction, 31*(2), 402–415. https://asjp.cerist.dz/en/article/277762

Al-Tarawneh, A. (2025). Optimizing multilingual projects: Exploring project management approaches in translation studies. In A. M. A. M. Al-Sartawi, M. Al-Okaily, A. A. Al-Qudah, & F. Shihadeh (Eds.), *From machine learning to artificial intelligence* (Studies in Systems, Decision and Control, Vol. 572, pp. 1709–1720). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-76011-2_128

Al-Tarawneh, A., & Al-Badawi, M. (2025). Exploring the synergy between translation practices and project management in multilingual project environments. In: A. Al-Sartawi & H. Ghura (Eds.), *Artificial intelligence, sustainable technologies, and business innovation: Opportunities and challenges of digital transformation* (Studies in Computational Intelligence, Vol. 1171, pp. 575–585). Springer. https://doi.org/10.1007/978-3-031-77925-1_50

Antúnez Brizuela, A. M. (2019). *El rol del gestor de proyectos en una empresa de traducción* (Trabajo de fin de máster). Universitat Oberta de Catalunya. https://openaccess.uoc.edu/handle/10609/91277

Bowker, L. & Buitrago Ciro, J. (2019). Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community. *The Journal of Specialised Translation, 35*, 233–234. https://doi.org/10.26034/cm.jostrans.2021.127

Cadwell, P., Castilho, S., O'Brien, S., & Mitchell, L. (2016). Human factors in machine translation and post-editing among institutional translators. *Translation Spaces,* 5(2), 222–243. https://doi.org/10.1075/ts.5.2.04cad

Chen, L. (2025). Examining cognitive load in human-machine collaborative translation: Insights from eye-tracking experiments of Chinese-English translation. *Frontiers in Psychology, 16*, 1570929. https://doi.org/10.3389/fpsyg.2025.1570929

Das, A. K. (2018). Translation and artificial intelligence: Where are we heading? *International Journal of Translation, 30*(1), 1–26.

Davitti, E. (2025). Technology for hybrid modalities. In: E. Davitti, T. Korybski, & S. Braun (Eds.), *The Routledge handbook of interpreting, technology and AI* (pp. 181–208). Routledge. https://doi.org/10.4324/9781003053248-15

Dinu, G., Mathur, P., Federico, M., & Al-Onaizan, Y. (2019). Training neural machine translation to apply terminology constraints. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 3063–3068). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1294

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Fuentes-Pérez, I., & Cano Fernández, C. (2025). La industria de la traducción y la interpretación más allá de la traducción y la interpretación: Las nuevas profesiones emergentes y el impacto de la tecnología en los nuevos perfiles profesionales. *Sendebar, 36*, 209–234. https://doi.org/10.30827/sendebar.v36.32396

Georgakopoulou, P. (2019). Template Files: The Holy Grail of Subtitling. *Journal of Audiovisual Translation, 2*(2), 137–160. https://doi.org/10.47476/jat.v2i2.84

Guerreiro, N. M., Voita, E., & Martins, A. (2023). Looking for a needle in a haystack: A comprehensive study of hallucinations in neural machine translation. In: A. Vlachos & I. Augenstein (Eds.), *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 1059–1075). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.eacl-main.75

Guerreiro, N. M., Voita, E., & Martins, A. F. T. (2023). Looking for a needle in a haystack: A comprehensive study of hallucinations in neural machine translation (arXiv:2208.05309v2) [cs.CL]. *arXiv*. https://doi.org/10.48550/arXiv.2208.05309

Hasler, E., de Gispert, A., Iglesias, G., & Byrne, B. (2018). Neural machine translation decoding with terminology constraints. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2: Short Papers (pp. 506–512). Association for Computational Linguistics. https://doi.org/10.18653/v1/N18-2081

Hiebl, B., & Gromann, D. (2023). Quality in human and machine translation: An interdisciplinary survey. In: *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (pp. 375–384). European Association for Machine Translation. https://aclanthology.org/2023.eamt-1.37.pdf?utm

International Organization for Standardization. (2017). ISO 18587:2017: Translation services—Post-editing of machine translation output—Requirements. https://www.iso.org/standard/62970.html

Kim, S., Sung, M., Lee, J., Lim, H., & Gimenez Perez, J. F. (2024). Efficient terminology integration for LLM-based translation in specialized domains (arXiv:2410.15690v1) [cs.CL]. *arXiv*. https://doi.org/10.48550/arXiv.2410.15690

Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation, 25*, 131–148. http://www.jostrans.org/issue25/art_koponen.pdf

Koponen, M., Aziz, W., Ramos, L., & Specia, L. (2012). Post-editing time as a measure of cognitive effort. In: S. O'Brien, M. Simard, & L. Specia

(Eds.), *Proceedings of the AMTA 2012 Workshop on Post-Editing Technology and Practice* (pp. 11–20). Association for Machine Translation in the Americas. https://aclanthology.org/2012.amta-wptp.2

Krüger, R. (2025). Implementing generative artificial intelligence technologies in language industry workflows: A competence perspective. *Lebende Sprachen, 70*(1), 11–38. https://doi.org/10.1515/les-2025-0016

Läubli, S., Sennrich, R., & Volk, M. (2018). Has machine translation achieved human parity? A case for document-level evaluation. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 4791–4796). Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-1512

Massion, F. (2021). Terminology in the age of artificial intelligence. *Journal of Translation Studies, 1*(1), 87–108. https://doi.org/10.3726/JTS012021.6

Mohamed, Y. A., Khanan, A., Bashir, M., Mohamed, A. H. H., Adiel, M. A. E., & Elsadig, M. A. (2024). The impact of artificial intelligence on language translation: A review. *IEEE Access, 12*, 25553–25579. https://doi.org/10.1109/ACCESS.2024.3366802

Moneus, A. M., & Sahari, Y. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon, 10*(6), e28106. https://doi.org/10.1016/j.heliyon.2024.e28106

Moorkens, J. (2017). Under pressure: translation in times of austerity. *Perspectives, 25*(3), 464–477. https://doi.org/10.1080/0907676X.2017.1285331

Moslem, Y., Haque, R., Kelleher, J., & Way, A. (2023, June). Adaptive machine translation with large language models. In: *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (pp. 227-237). https://doi.org/10.48550/arXiv.2301.13294

Moslem, Y., Romani, G., Molaei, M., Kelleher, J. D., Haque, R., & Way, A. (2023). Domain terminology integration into machine translation: Leveraging large language models. In: *Proceedings of the Eighth Conference on Machine Translation* (pp. 902–911). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.wmt-1.82

O'Brien, S. (2021). Post-editing. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 5, pp. 177–184). John Benjamins. https://doi.org/10.1075/hts.5.pos3

O'Brien, S. (2022). How to deal with errors in machine translation: post-editing. In: D. Kenny (Ed.), *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence* (pp. 105–120). https://doi.org/10.5281/zenodo.6759982

O'Brien, S., & Mellinger, C. D. (2025). Research methods in translation and interpreting studies. In: C. A. Chapelle, B. Baer & C. Mellinger (Eds.), *The encyclopedia of applied linguistics: Translation and interpreting. Wiley*. https://doi.org/10.1002/9781405198431.wbeal20735

Pinnis, M. (2016). Towards hybrid neural machine translation for English–Latvian. In: I. Skadiņa & R. Rozīte (Eds.), *Proceedings of the 7th International Conference on Human Language Technologies – The Baltic Perspective*

(Baltic HLT 2016) (pp. 84–91). IOS Press. https://doi.org/10.3233/978-1-61499-701-6-84

Pinnis, M., Kalniņš, R., Skadins, R., & Skadiņa, I. (2016). What can we really learn from post-editing? In: *Proceedings of AMTA 2016*, Vol. 2: MT Users' Track (pp. 86–91). The Association for Machine Translation in the Americas. https://aclanthology.org/2016.amta-users.8.pdf

Pinnis, M., Krišlauks, R., Deksne, D., & Miks, T. (2017). Neural machine translation for morphologically rich languages with improved sub-word units and synthetic data. In: K. Ekštein & V. Matoušek (Eds.), *Text, Speech, and Dialogue* (pp. 237–245). Springer. https://doi.org/10.1007/978-3-319-64206-2_27

Qin, L., Chen, Q., Zhou, Y., Chen, Z., Li, Y., Liao, L., Li, M., Che, W., & Yu, P. S. (2024). A survey of multilingual large language models. *Patterns, 5*(10), Article 101118. https://doi.org/10.1016/j.patter.2024.101118

Rico Pérez, C. (2024). Re-thinking machine translation post-editing guidelines. *The Journal of Specialised Translation, 41*, 26–47. https://doi.org/10.26034/cm.jostrans.2024.4696

Risku, H., & Windhager, F. (2015). Extended translation: A sociocognitive research agenda. In M. Ehrensberger-Dow, S. Göpferich, & S. O'Brien (Eds.), *Interdisciplinarity in translation and interpreting process research* (pp. 35–47). John Benjamins. https://doi.org/10.1075/bct.72.04ris

Risku, H., Milošević, J., & Rogl, R. (2017). Final report of the research project "Extended translation: Socio-cognitive translation processes in the workplace" [Final project report]. *Centre for Translation Studies, University of Vienna.* https://socotrans.univie.ac.at/fileadmin/user_upload/p_socotrans/20190522_Summary_ExTra_Final_report.pdf

Risku, H., Milošević, J., & Rogl, R. (2019). Changes in the field of translation project management: Findings of a longitudinal ethnographic study (Working paper). *Centre for Translation Studies, University of Vienna*. https://socotrans.univie.ac.at/fileadmin/user_upload/p_socotrans/20190613_Working_Paper_Changes_Risku_Milosevic_Rogl.pdf

Rogl, R., & Risku, H. (2024). Cognitive artefacts and boundary objects: On the changing role of tools in translation project management. In: M. Winters, S. Deane-Cox, & U. Böser (Eds.), *Translation, interpreting and technological change: Innovations in research, practice and training* (pp. 13–36). Bloomsbury Publishing.

Saldanha, G., & O'Brien, S. (2013). *Research methodologies in translation studies* (1st ed.). Routledge. https://doi.org/10.4324/9781315760100

Sánchez-Gijón, P., Moorkens, J., & Way, A. (2019). Post-editing neural machine translation versus translation memory segments. *Machine Translation, 33*(1/2), 31–59. http://www.jstor.org/stable/45222461

Shahmerdanova, R. (2025). Artificial intelligence in translation: Challenges and opportunities. *Acta Globalis Humanitatis et Linguarum, 2*(1), 62–70. https://doi.org/10.69760/aghel.02500108

Strandvik, I. (2025). Translation quality and the role of specifications: How standards can help the translation sector today. *Across Languages and Cultures, 26*(S), 5–28. https://doi.org/10.1556/084.2025.01057

Tarawneh, A. (2025). Efficient project management strategies for translation projects: Maximizing quality and timeliness in multilingual content creation. In: A. Al-Sartawi & H. Ghura (Eds.), *Artificial intelligence, sustainable technologies, and business innovation: Opportunities and challenges of digital transformation (Studies in Computational Intelligence, Vol. 1171*, pp. 755–764). Springer. https://doi.org/10.1007/978-3-031-77925-1_66

TAUS. (2023). *Post-editing trends and pricing in the AI-driven translation industry*. TAUS Report.

Van der Meer, J., & Federmann, C. (2020). *Advances and challenges in neural machine translation: Industry insights.* TAUS Review.

Vieira, L. N. (2014). Indices of cognitive effort in machine translation post-editing. *Machine Translation, 28*(3–4), 187–216. https://doi.org/10.1007/s10590-014-9156-x

Wang, Y., & Jalalian Daghigh, A. (2023). Cognitive effort in human translation and machine translation post-editing processes: A holistic and phased view. FORUM. *Revue internationale d'interprétation et de traduction / International Journal of Interpretation and Translation, 21*(1), 139–162. https://doi.org/10.1075/forum.22009.wan

Xu, W., Agrawal, S., Briakou, E., Martindale, M. J., & Carpuat, M. (2023). Understanding and detecting hallucinations in neural machine translation via model introspection. *Transactions of the Association for Computational Linguistics, 11*, 546-564. https://doi.org/10.48550/arXiv.2301.07779

Anexos (A-B-C)

Anexo A. Ejemplo ilustrativo de glosario terminológico gobernado para flujos híbridos (TA/IA + posedición)

A.1 Propósito del glosario en entornos híbridos

En proyectos de traducción mediados por traducción automática e inteligencia artificial (TA/IA), un glosario terminológico no funciona únicamente como una lista de equivalencias. En estos entornos, el glosario opera como un mecanismo de gobernanza terminológica, ya que establece qué términos deben utilizarse, cuáles deben evitarse y bajo qué condiciones, con el fin de asegurar consistencia, trazabilidad y cumplimiento de requisitos del cliente a lo largo del flujo (pretraducción, posedición, aseguramiento de calidad y validación final). Por ello, cada entrada puede incorporar metadatos operativos (estado de aprobación, responsable, versión, prioridad, dominio, notas de uso y restricciones), lo que permite integrar el glosario en herramientas CAT y en reglas automatizadas de QA.

Siglas usadas (aclaración breve):
PM = *Project Manager* (gestor/a de proyectos) · **QA** = *Quality Assurance* (aseguramiento de calidad) · **CAT tools** = *Computer-Assisted Translation tools* (herramientas de traducción asistida) · **TA/IA** = traducción automática/inteligencia artificial.

A.2 Glosario ilustrativo (ES → EN) con control terminológico y restricciones operativas
Proyecto: Plataforma de e-commerce (UI + ayuda al usuario)
Locales: ES (rioplatense) → EN (US)
Versión del glosario: v1.0 | **Estado global:** Aprobado para producción
Responsable: PM / Terminología

A.3 Uso operativo del glosario en flujos híbridos (resumen)
1. **Pretraducción (MT/IA):** el glosario guía la terminología preferida.
2. **Posedición:** se corrigen variantes "no usar".
   QA automatizado: reglas detectan incumplimientos terminológicos.
3. **Revisión especializada:** validación en segmentos críticos.
4. **Actualización/versionado:** cambios registrados con responsable, fecha y justificación.

ANEXO B. Ejemplo ilustrativo de glosario técnico-regulatorio (médico/jurídico) gobernado para flujos híbridos (TA/IA + posedición)

B.1 Propósito del glosario en entornos regulados
En proyectos médico-regulatorios y jurídicos, el glosario terminológico opera como un instrumento de control de riesgo. Su función es garantizar consistencia y precisión en términos críticos (p. ej., seguridad del paciente, cumplimiento normativo, validez contractual), evitando sustituciones por sinónimos, paráfrasis o equivalentes no estándares que pueden aparecer en salidas automatizadas.

Siglas usadas (aclaración breve):
SME = *Subject Matter Expert* (especialista del dominio: médico/regulatorio/jurídico) · **IFU** = *Instructions for Use* (instrucciones de uso) · **PMS** = *Post-Market Surveillance* (vigilancia poscomercialización) · **QA** = aseguramiento de calidad · **PM** = gestor/a de proyectos.

B.2 Glosario ilustrativo (ES → EN) con criticidad y reglas de uso
Proyecto: Dispositivo médico + documentación contractual UE/US
Locales: ES (LatAm) → EN (US)
Versión: v1.0 | **Estado:** Aprobado para producción
Responsables: PM + SME médico + asesor legal
Criticidad: Crítico (cumplimiento/seguridad/validez legal) **Mayor** (precisión técnica/consistencia) · **Menor** (preferencia secundaria)

ID	Término fuente (ES)	Equivalente preferido / no usar	Dominio, criticidad y estado	Contexto y regla de uso
001	evento adverso	Preferido: *adverse event*. No usar: *side effect*.	Médico; crítico; aprobado	"Notificar evento adverso". *Side effect* se restringe a fármacos; usar *adverse event*.
002	incidente grave	Preferido: *serious incident*. No usar: *severe incident*.	Regulatorio; crítico; aprobado	"Incidente grave debe reportarse". Evitar *severe*; término estandarizado.
003	contraindicación	Preferido: *contraindication*. No usar: *not recommended*.	Médico; crítico; aprobado	"Contraindicaciones". Paráfrasis ambigua; usar término técnico.
004	advertencia	Preferido: *warning*. No usar: *caution*.	Médico; mayor; aprobado	"Advertencias". *Warning* indica riesgo serio; *caution*, riesgo moderado.
005	precaución	Preferido: *caution*. No usar: *warning*.	Médico; mayor; aprobado	"Precauciones". Mantener distinción *warning/caution*.
006	instrucciones de uso	Preferido: *instructions for use* (IFU). No usar: *user guide*.	Regulatorio; crítico; aprobado	"Consulte las IFU". IFU es obligatorio en documentación regulatoria.
007	manual del usuario	Preferido: *user manual*. No usar: *instructions for use*.	Técnico; mayor; aprobado	"Manual del usuario". No confundir con IFU.
008	marcado CE	Preferido: *CE marking*. No usar: *CE mark*.	Regulatorio; crítico; aprobado	"Con marcado CE". Preferir *CE marking*.
009	evaluación clínica	Preferido: *clinical evaluation*. No usar: *medical evaluation*.	Regulatorio; crítico; aprobado	"Informe de evaluación clínica". Término estándar.
010	vigilancia poscomercialización	Preferido: *post-market surveillance* (PMS). No usar: *after-sales monitoring*.	Regulatorio; crítico; aprobado	"Plan PMS". Evitar paráfrasis; mantener PMS.
011	cumplimiento	Preferido: *compliance*. No usar: *conformity*.	Regulatorio; mayor; aprobado	"Cumplimiento normativo". Usar *compliance* para cumplimiento de normas.
012	auditoría	Preferido: *audit*. No usar: *inspection*.	Legal/QA; mayor; aprobado	"Auditoría interna". *Inspection* se reserva para autoridades/regulador.
013	inspección	Preferido: *inspection*. No usar: *audit*.	Legal/QA; mayor; aprobado	"Inspección del organismo". Mantener distinción.

014	exención de responsabilidad	Preferido: *disclaimer*. No usar: *responsibility exemption*.	Jurídico; crítico; aprobado	"Exención de responsabilidad". Usar término estándar.
015	incumplimiento	Preferido: *breach*. No usar: *noncompliance*.	Jurídico; crítico; aprobado	"Incumplimiento del contrato". En contratos, usar *breach*.
016	rescisión	Preferido: *termination*. No usar: *cancellation*.	Jurídico; mayor; aprobado	"Rescisión del contrato". Preferir *termination*.
017	fuerza mayor	Preferido: *force majeure*. No usar: *act of God*.	Jurídico; mayor; aprobado	"Cláusula de fuerza mayor". Preferir término estándar.

B.3 Metadatos de gobernanza recomendados (por entrada)
Owner: PM (gestor/a)
Validación SME: revisión del especialista del dominio (médico/legal)
Fuente: IFU aprobadas, normativa del cliente, cláusulas modelo, documentación oficial
Criticidad: Crítico / Mayor / Menor
Regla QA: error crítico / mayor / advertencia
Versionado: v1.0, v1.1… + fecha, responsable, justificación

Notas: obligatoriedad, contexto permitido/restringido, variantes prohibidas

Anexo C. Reglas de QA y tratamiento de desviaciones según criticidad terminológica
Este cuadro muestra cómo la gobernanza terminológica se operacionaliza mediante reglas de QA en un flujo híbrido, diferenciando umbrales de tolerancia y responsabilidades según el impacto potencial de cada desviación.

Nivel de criticidad	Definición operativa	Regla de QA (detección)	Acción requerida	Responsable	Criterio de aceptación
Crítico	La desviación puede afectar cumplimiento regulatorio, seguridad o validez contractual	Detección automática + lista "No usar" + verificación manual en segmentos sensibles	*Corrección inmediata obligatoria* + registro de incidencia + revisión del flujo si se repite	PM + SME + revisor senior	0 desviaciones permitidas

Mayor	Afecta precisión y consistencia técnica, sin impacto directo inmediato en compliance	Detección automática + muestreo dirigido + reglas de consistencia	Corrección obligatoria antes de entrega final + nota en QA report	PM + revisor técnico	Desviación tolerable solo si está *justificada y aprobada*
Menor	Preferencia estilística o consistencia secundaria	Alerta QA (advertencia) + revisión superficial	Corrección recomendada si mejora consistencia; no bloquea entrega	Poseditor + PM	Puede aceptarse si no altera significado ni consistencia global