

UNIVERSIDAD TECNOLÓGICA DE BOLÍVAR

INTELIGENCIA ARTIFICIAL

Taller sobre la base de conocimientos: Machine translation

Martinez Palmieri Kevin Esteban, T00061428

Martínez Velandia Jesús Antonio, T00061258

Prof. Edwin Puertas

21 de octubre de 2023

Resumen

Esta ontología es una herramienta esencial para comprender y analizar los sistemas de traducción automática, centrándose en mejorar la eficacia y la precisión de las traducciones. Esta ontología proporciona una base sólida para optimizar la interacción entre lenguas y culturas, facilitando la comprensión y la colaboración global, a medida que la comunicación global sigue siendo un imperativo en la era digital.

1. Introducción

La creación de una ontología que represente un sistema de traducción automática y sus componentes subyacentes es el tema principal de este estudio. La esencia del proceso de traducción automática, desde la entrada de texto en un idioma específico hasta la creación de una traducción en otro idioma, se captura en la ontología.

Hemos definido clases que representan entidades clave, como el “Sistema de traducción”, el “Motor de traducción”, la interfaz de usuario, el “texto de entrada” y el “texto traducido”, Además, hemos creado características y relaciones que representan las interacciones entre estas entidades. Estas características incluyen la acción de traducción en sí, la presentación de la traducción al usuario y la calidad de la traducción resultante.

La ontología no solo tiene como objetivo comprender e ilustrar los aspectos técnicos de la traducción automática, sino también proporcionar una base para el análisis y la mejora de la precisión de las traducciones. En un mundo donde la información fluye constantemente a través de las fronteras lingüísticas, la calidad de las traducciones es crucial. Por lo tanto, el trabajo examina y mide la calidad de las traducciones utilizando datos de referencia y comentarios de los usuarios.

2. Objetivo general

El objetivo principal de la base de conocimientos es modelar y representar un sistema de traducción automática como Google Translate, junto con sus componentes y relaciones. La base de conocimientos tiene como objetivo recopilar los datos cruciales que definen este sistema, así como sus interacciones con los usuarios, los textos de entrada y los resultados de traducción.

3. Marco teórico

Se procederá a dar una breve descripción de los conceptos que usaron al momento de construir la ontología, lo anterior, con la finalidad de que todos los lectores puedan comprender de manera clara la información que proporcionamos:

3.1. Protégé

La arquitectura de plug-ins de Protégé puede adaptarse para crear aplicaciones basadas en ontologías tanto sencillas como complejas. Los desarrolladores pueden integrar la salida de Protégé con sistemas de reglas u otros solucionadores de problemas para construir una amplia gama de sistemas inteligentes. Y lo que es más importante, el equipo de Stanford y la amplia comunidad de Protégé están aquí para ayudar.

3.2. Ontología

Las ontologías, en inteligencia artificial, son estructuras conceptuales creadas por expertos para representar el conocimiento de una manera explícita y formal y que pueda ser procesable por las máquinas.

3.3. Machine Translation

Un servicio de traducción en línea llamado Machine Translation (MT) traduce texto de un idioma en otro mediante el uso de inteligencia artificial y algoritmos sofisticados. Tiene más de 100 idiomas compatibles y otras funciones como detección de idiomas y comparación de texto.

3.4. Idioma

Sistema de comunicación propio de una sociedad humana, ya sea verbal (lengua oral y gráfica) o gestual (lengua signada).

3.5. Texto

Un texto es una unidad de sentido formada por signos codificados en un sistema de escritura.

3.6. Modelo Traducción

El modelo de traducción es un enfoque para enseñar idiomas que se basa en el análisis minucioso de las reglas gramaticales y sus excepciones. Luego, estos conocimientos se aplican a la traducción de oraciones y textos de la lengua de destino a la propia y a la inversa. En otras palabras, este enfoque se centra en la traducción como método para aprender un idioma.

4. Desarrollo

Para la construcción de esta ontología se usó el documento y las diapositivas referentes a la creación de la ontología propuestos en clase, cabe resaltar que para la resolución de la ontología se implementó el software de "**Protégé**" que es uno de los software más utilizados para la creación de este tipo de prácticas.

4.1. Determinar el dominio y alcance de la ontología

Alcance específico

La base de conocimientos se centra en la representación de los sistemas de traducción automática, incluida su configuración, componentes, procesos y resultados. No incluye información detallada sobre la tecnología utilizada en los motores de traducción o sistemas particulares como Google Translate. Se enfoca en la modelización conceptual de las relaciones y componentes clave.

Dominio específico

Todos los aspectos, conexiones y conceptos relacionados con los sistemas de traducción automática y su funcionamiento están incluidos en el dominio "Traducción automática". El sistema de traducción propio, los motores de traducción, las interfaces de usuario, los textos de entrada y los textos traducidos, las lenguas de origen y de destino y la calidad de la traducción son todos componentes del paquete.

4.2. Considerar la reutilización de ontologías existentes

No se reutilizó una ontología existente, se optó por crear una desde 0.

4.3. Enumerar términos importantes para la ontología

Los términos más importantes que se consideraron para la construcción de esta ontología fueron:

- Sistema de Traducción.
- Interfaz de Usuario.
- Motor de Traducción.
- Texto de Entrada.
- Texto Traducido.

4.4. Definir las clases y la jerarquía de clases

Para hacer la antología se necesita de clases y jerarquías que se definirán a continuación:

- **Sistema de Traducción:** Representa una instancia específica de un sistema de traducción automática, como "Google Translate."
- **Interfaz de Usuario:** Refleja la interfaz a través de la cual los usuarios interactúan con el sistema de traducción.

- **Motor de Traducción:** Representa el motor principal que realiza las traducciones.
- **Texto de Entrada:** Incluye los textos que los usuarios desean traducir. Debe incluir detalles del idioma de origen.
- **Texto Traducido:** Representa los resultados de las traducciones. Debe incluir detalles del idioma de destino y la calidad de la traducción.

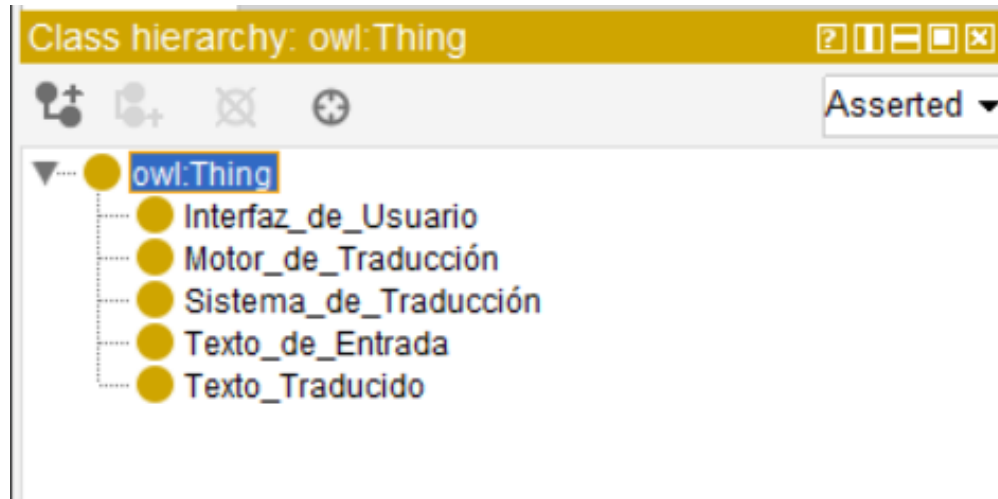


Figura 4.1: Visualización de las entidades por medio del software elegido.

No hay jerarquía en ninguna de las clases.

4.5. Definir las propiedades de las clases (slots)

Propiedades y Relaciones Relevantes:

Las propiedades y relaciones relevantes que se consideraron fueron las siguientes:

- **Traduce Texto:** Establece la relación entre el Sistema de Traducción y el Motor de Traducción, indicando que el sistema utiliza un motor específico para realizar traducciones.
- **Muestra Traducción:** Representa la conexión entre el Sistema de Traducción y la Interfaz de Usuario, demostrando cómo se presenta la traducción al usuario.

- **Tiene Texto de Entrada:** Conecta el Sistema de Traducción con el Texto de Entrada, especificando qué texto se traducirá.
- **Genera Texto Traducido:** Establece la relación entre el Motor de Traducción y el Texto Traducido, indicando que el motor produce la traducción.
- **Usa Lenguaje de Origen:** Asocia el Texto de Entrada con su idioma de origen, lo que ayuda a definir la traducción.
- **Usa Lenguaje de Destino:** Vincula el Texto Traducido con el idioma al que se tradujo, permitiendo el seguimiento de la traducción.
- **Calidad de Traducción:** Describe la calidad de la traducción en el Texto Traducido, ya sea en forma numérica o categorías de calidad.

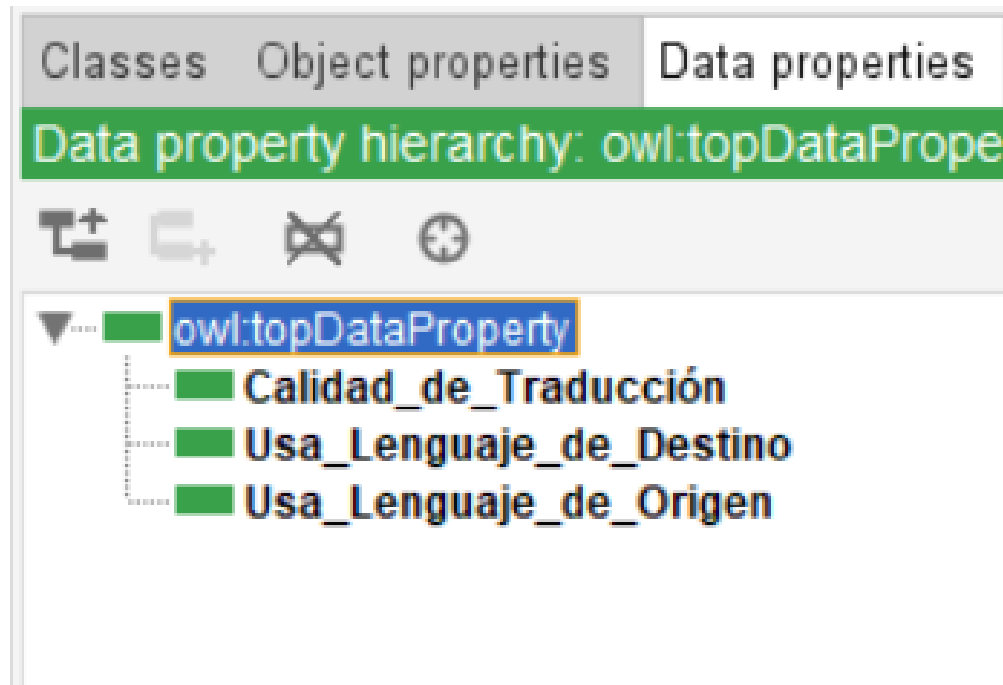


Figura 4.2: Visualización de las propiedades de los datos por medio del software elegido.

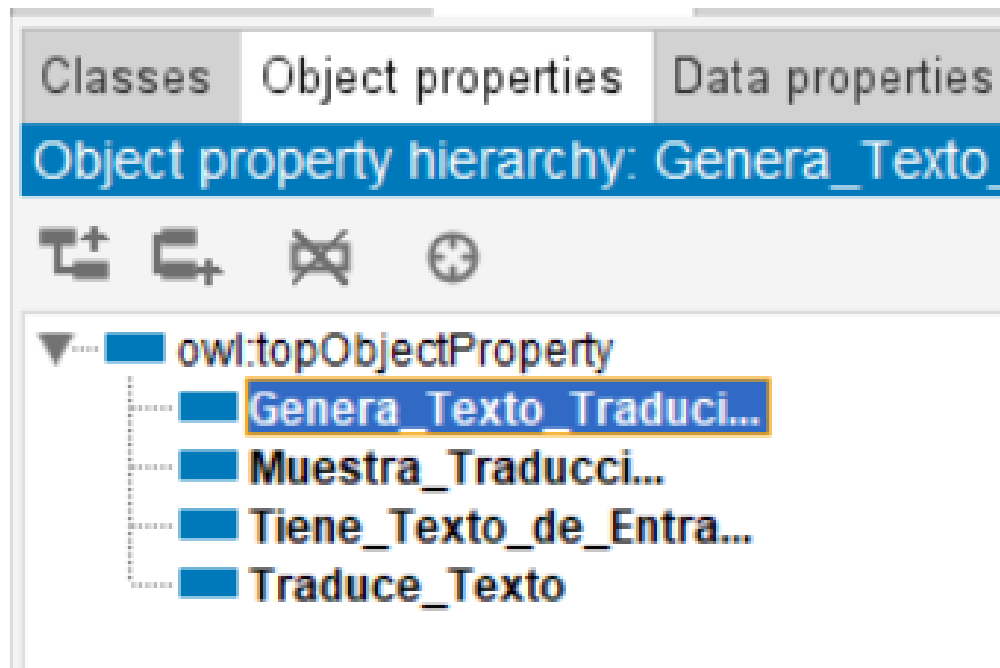


Figura 4.3: Visualización de las propiedades de relación de entidades por medio del software elegido.

4.6. Definir las instancias

las instancias que se tomaron en cuenta fueron: las anteriores instancias se usan más que

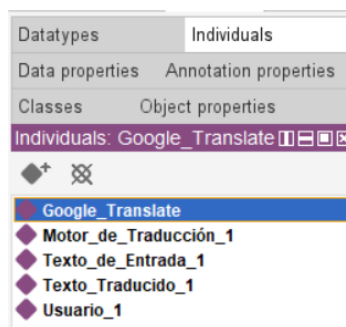


Figura 4.4: Visualización de las instancias por medio del software elegido.

nada para probar que todo esté marchando como debe ser, en este caso solo se coloca una prueba desde uno de los infinitos casos hipotéticos que existen con la finalidad de demostrar la eficacia de la ontología.

4.7. Visualización de los grafos y ejecución de la ontología

Para la visualización de los grafos se implementará la herramienta “**OntoGraf**” que incluye el software, se escogió, está por encima de las otras porque es muy fácil e intuitiva de utilizar, el resultado que nos arrojó fue el siguiente:

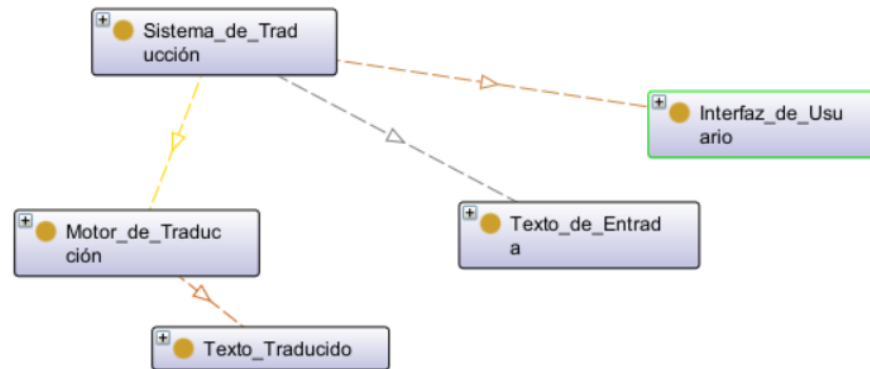


Figura 4.5: Visualización de los grafos por medio del software elegido.

Para la visualización de la ejecución de la ontología se procedió a ejecutarse por medio de “**HermiT 1.4.3.456**” que uno de los compiladores por excelencia que cuenta el software, dicha herramienta nos arroja el siguiente resultado:

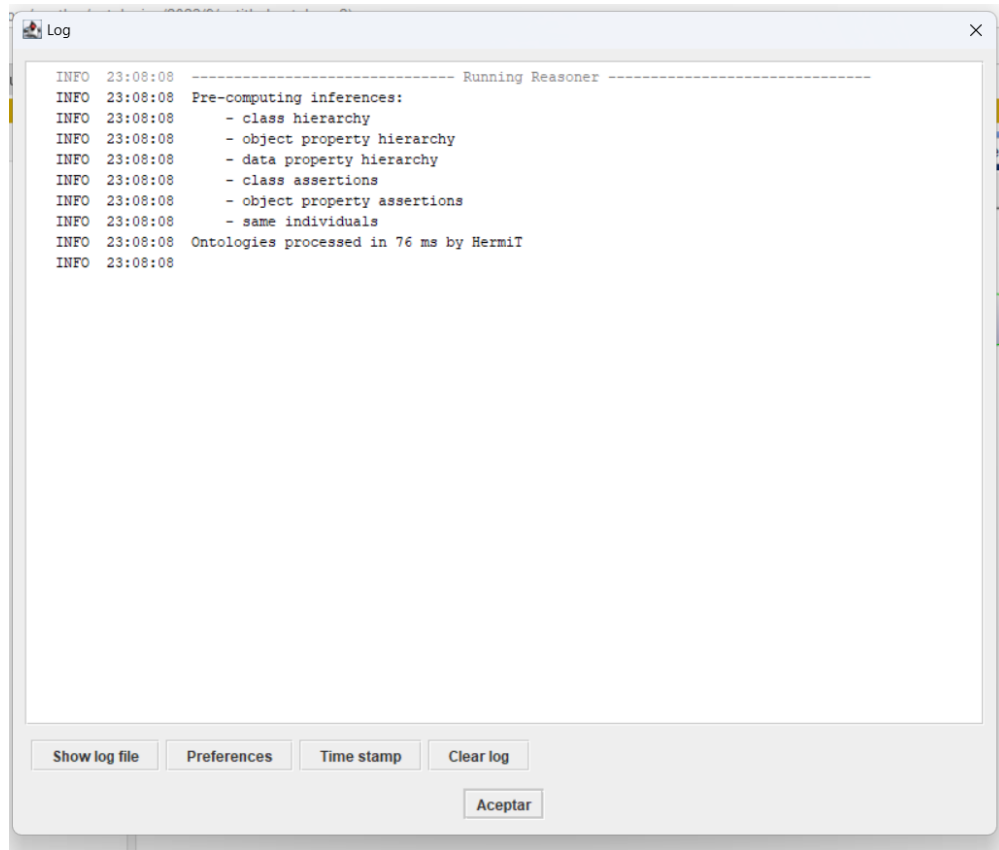


Figura 4.6: Visualización del resultado de la ontología por medio del software elegido.

5. Conclusiones

En conclusión, la elaboración de la ontología nos ayudó a comprender mejor el funcionamiento del agente inteligente propuesto en clase, viendo su funcionamiento con respecto a unos datos de pruebas proporcionados por los desarrolladores, con la finalidad de poder visualizar de mejor manera como son procesados los datos.

Para la correcta visualización de la ontología se le recomienda al lector que ingrese en GitHub y descargue el archivo .owl, para que posteriormente los pueda abrir por medio del software de “Protégé” y ejecutar el código, aunque si solo se requiere ver, se puede visualizar directamente en el GitHub.