

Prototipos para la exploración de términos Kichwa representados como Grafo de Conocimiento basado en Tecnologías de la Web Semántica

Nelson Piedra, Klever Fernando Sarango Saca

Departamento de Ciencias de la Computación y Electrónica

Universidad Técnica Particular de Loja

Loja, Ecuador

nopiedra@utpl.edu.ec, kfsarango1@utpl.edu.ec

Resumen— Las interfaces de conversación o chatbots están alterando los canales tradicionales de interacción entre organizaciones y usuarios. Las organizaciones buscan ofrecer servicios de soporte 24/7. Los chatbots permiten simular una conversación inteligente con los usuarios, ya sea a través de mensajes de texto y/o audio. Son utilizados para ofrecer información bajo demanda, facilitar orientación, resolver dudas, brindar soporte, y muchos otros servicios funcionales más. Las áreas de servicios de las empresas los utilizan cada vez más para interactuar con sus clientes. Los chatbots más avanzados se alimentan con base de conocimiento semántica, algoritmos de aprendizaje automático y procesamiento de lenguaje natural, que permiten “comprender” solicitudes complejas, personalizar respuestas y mejorar las interacciones a lo largo del tiempo.

Este documento informa sobre el desarrollo de una interfaz conversacional para responder a las demandas de los estudiantes cuando buscan información académica sobre cursos abiertos. La interfaz conversacional desarrollada proporcionará información personalizada y orientación a los usuarios, dentro de las acciones de diccionario Kichwa – Español - Inglés. Los prototipos están desarrollados para los servicios Telegram y Twitter. La información usada en la interfaz conversacional está modelada como un grafo de conocimiento (KG) implementado con tecnologías y estándares de la Web Semántica (SW). La intención de los KG es organizar la información con formatos procesables directamente por máquinas y enlazados semánticamente, de manera que los usuarios puedan navegar de una manera usable a través del KG creado. Se diseñó un KG con entradas de lenguaje Kichwa y su significado en Castellano e Inglés. Se construyó una base de conocimiento usando el enfoque de datos enlazados. Las entidades, propiedades y relaciones que forman parte del KG producido están basadas en los procesos de extracción y estructuración desde diccionario creados en Ecuador.

Keywords; kichwa; aprendizaje, web, bot;

I. CONTEXTO

De acuerdo a [1] Usamos “qichwa” para lo general y “kichwa” para la variante ecuatoriana, según la tradición de la Andinística. El idioma Kichwa cuyo significado es “Lengua de hombres”, es una lengua Antigua hablada por los incas, que se fue propagando en los pueblos conquistados por los mismos. Es una de las 14 lenguas ancestrales que se hablan en nuestro país, reconocida en la constitución ecuatoriana (ver Figura 1). Tiene una forma muy particular en su escritura en relación al español, durante el II Encuentro del 21 al 23 de julio de 1998, en Tabacundo (Pichincha), se ratificó el sistema grafémico de 1997 con ligerísimas variaciones y su uso [1]. Permite que a una sola palabra se unan varios conceptos, por lo que se considera como una lengua aglutinante, es decir que en muchos de los casos se forma con la unión de varios afijos o palabras.

En noviembre del 2003, en Ecuador la Dirección Nacional de Educación Intercultural Bilingüe (DINEIB) creó la Academia de la Lengua Kichwa/ALKI, esto es la “Kichwa Amawta Kamachik” –KAMAK-. La KAMAK se encargó del estudio, investigación, vigilancia, decretos, interrelaciones y proyectos para el desarrollo del kichwa tanto al interior del Ecuador así como a nivel de los demás países de habla qichwa. Con esta finalidad, KAMAK se reunió del 22 al 26 de marzo del 2004, en Riobamba para legitimar una vez más el sistema de 1997. Un hecho destacable es es la mención de la estructuración del alfabeto en base a criterios del Alfabeto Fonético Internacional (AFI) y con miras a uno de alcance “Panandino” [1].



Fig. 1. Pueblos y Nacionalidades Indígenas en Ecuador [1]

La Subsecretaría de Educación para el Diálogo Intercultural (SEPDI), con la finalidad de elaborar un Diccionario Escolar Intercultural Bilingüe kichwa-castellano-kichwa como texto en los procesos de Educación Intercultural Bilingüe (EIB), con la participación de varios delegados de la DINEIB, las Direcciones Provinciales de Educación Intercultural Bilingüe (DIPEIBs), Centros Educativos, Instituciones relacionadas con EIB, Universidades, y otras organizaciones, ejecutó varios talleres sobre lexicografía, historia, lengua y cultura qichwas y

metodología de enseñanza de lenguas [1]. Además, se ejecutaron varias actividades sobre revisión, modificación, actualización y conocimiento de diccionarios kichwas creados a lo largo de su historia.

II. OBJETIVOS

A. General

- 1) Desarrollar un Knowledge Base para el idioma Kichwa, utilizando un enfoque Linked Data.

B. Específicos

- 1) Colectar y transformar a información estructurada diccionarios relacionadas con el idioma Kichwa.
- 2) Diseñar una arquitectura para la extracción de entidades, relaciones semánticas, enlazado y construcción del grafo de conocimiento.
- 3) Implementar las acciones de extracción y enlazado de entidades y construir el grafo de conocimiento.
- 4) Desarrollar prototipo de exploración del grafo basado en un Chatbot.

III. DESCRIPCIÓN DE APLICACIONES INFORMÁTICAS RELACIONADAS CON KICHWA

A continuación se presentan las ideas de un conjunto de trabajos analizados para llevar a cabo la definición, modelado y desarrollo del presente trabajo. Se hablará brevemente de cada uno, dando a conocer los puntos más sobresalientes que han servido de referencia para el presente trabajo.

A. *Aplicación Relacionada 1 - Traductor Español Kichwa*

1) *Descripción:* La Carrera de Sistemas y Computación de la UNACH desarrolló el traductor denominado “Mashi - Kutipak” que significa “Amigo Traductor”, este traductor fue desarrollado para traducir de español al idioma Kichwa.

2) *Objetivo:* Con el cual se pretende que la información que el usuario requiera obtener en Kichwa sea confiable, sencilla de obtener.[2]

3) *Desarrollo:* Se basa específicamente en dos funciones principales:

- Proceso de traducción de palabras(Diccionario): Este proceso realiza la traducción automática basado en el modelo de transferencia, el texto original se analiza primero morfológica y sintácticamente, obteniendo como resultado una representación sintáctica superficial.[2]
- Proceso de traducción de contenidos: Este proceso se realiza cuando el usuario desea obtener traducciones de más de una palabra el texto es analizado morfológica y sintácticamente buscando lo requerido.[2]

4) *Orientación:* Es una aplicación educativa dirigida especialmente a los estudiantes, docentes y personal administrativo, como herramienta de apoyo para poder impulsar la Interculturalidad en las Aulas de la Institución.[2]

5) *Herramientas:* Para el desarrollo del aplicativo se utilizó las siguientes herramientas: Netbeans Ide Version 8.02; Mysql (Xampp V 1.8); Apache Tomcat (V 7.0); Protégé (V 5.0.0)

B. Aplicación Relacionada 2 - Reconocimiento de voz

1)*Descripción:* Es una aplicación informática de escritorio con reconocimiento de voz, contiene una agradable interfaz gráfica, de acuerdo a las necesidades de los usuarios, de tal forma que sea más fácil de utilizar y entender.[3]

2)*Objetivo:* Promover el aprendizaje básico del idioma Kichwa.

3)*Desarrollo:* El software consta de 5 módulos, donde se va a tratar diversos temas como: Saludos; Cuerpo Humano; Preguntas de información; Animales; Alimentos; Colores; Miembros de la familia, etc. Además de su implementación de traducción por voz, donde recibe como entrada la voz del usuario, para posteriormente ser traducida.

4)*Orientación:* El uso de este software educativo permite a los estudiantes de primer nivel del instituto de soluciones académicas alto nivel en Quito, estar más motivados por el tema de estudios y avanzar en el tema a su propio ritmo, según sus posibilidades y necesidades. [3] Con el desarrollo de este software se mejora la comprensión entre las personas que constituyen la sociedad facilitando la comunicación y comprensión del idioma Kichwa, este será los resultados a corto plazo, el no olvidar nuestro idioma nativo y consérvalo para las futuras generaciones. [3].

5)*Herramientas:* No se dan a conocer las herramientas a utilizar.

C. Aplicación Relacionada 3 - Realidad virtual y aumentada

1)*Descripción:* Se trata de una aplicación móvil denominada “Yachani Rayku Pakta” en español “Aprendo por igual”, que utiliza la realidad virtual y aumentada.

2)*Objetivo:* Desarrollar la aplicación como medio de aprendizaje del idioma Kichwa.

3)*Desarrollo:* El aplicativo presenta un libro físico de marcadores, considerando que, al acercar la cámara del dispositivo móvil a las imágenes del libro, se generan imágenes 3D del objeto y la respectiva traducción en el idioma Kichwa.[4]

4)*Orientación:* Es una aplicación educativa dirigida para niños entre 6 a 7 años. Tras el uso de la aplicación, los docentes han constatado que existe un incremento del interés del Kichwa por parte de los alumnos de la escuela de Pueblo Viejo del cantón Alausí, donde se está utilizando esta app.[4]

5)*Herramientas:* Para el desarrollo del aplicativo se utilizó las siguientes herramientas: Unity 5.4.0; SKD de Vuforia; Librería de google VR; Metodología XP

D. Aplicación Relacionada 4 - Enseñanza multimedia del Kichwa

1)*Descripción:* Se trata de una aplicación web denominada “Kichwa Multimedia” caracterizada por la implementación de una base de datos no relacional.

2)*Objetivo:* Describir conceptos, el idioma Kichwa y la metodología de enseñanza de idiomas.

3)*Desarrollo:* El aplicativo consta de los siguientes módulos:

- Módulo para la gestión del contenido Kichwa Multimedia.
- Módulo de diccionario Kichwa-Español-Ingles
- Módulo de juegos en Kichwa.
- Módulo de cursos y ejemplo de evaluación.
- Módulo de administración de usuarios.

4)*Orientación:* La correcta ejecución de este proyecto es un gran aporte al proceso de revitalización lingüística identitaria del Kichwa, el Kichwa Sisariy, beneficiando directamente a todas las personas que quieran o necesiten aprender Kichwa y los Kichwa parlantes que quieran mejorar su escritura y lectura.[5]

5) *Herramientas:* Para el desarrollo del aplicativo se utilizó las siguientes herramientas: Mongo DB; Java; Framework JSF; Apache Tomcat

E. Aplicación Relacionada 5 - Diccionario y traductor Español-Kichwa

1)*Descripción:* Es una aplicación que fue diseñada por un tesista de la Universidad Nacional de Loja, esta aplicación implementa un diccionario traductor Kichwa-Español.

2)*Objetivo:* Implementación de un sistema de apoyo pedagógico en la enseñanza del idioma Kichwa.

3)*Desarrollo:* La aplicación permite al estudiante consultar la escritura de las palabras del idioma Kichwa con su respectivo significado; también puede escuchar la correcta pronunciación de la misma; conjuntamente permite la traducción de frases desde el español.[6]

4)*Orientación:* Esta aplicación fue creada para los niños de educación básica de la escuela José Antonio Enrique de la ciudad de Loja.

5) *Herramientas:* Para el desarrollo del aplicativo se utilizó las siguientes herramientas:

- Iconix, como metodología de desarrollo; Java; Java Media Framework; Sistema Multimedia; MYSQL

F. Aplicación Relacionada 6 - Páginas web

Existen algunos sitios web específicamente de tipo gubernamentales, que implementan el idioma Kichwa en su sitio web, como una opción de multilinguaje, el mismo que maneja contenido traducido en Kichwa, a partir del contenido español normal. Estas páginas pertenecen a las siguientes entidades: Policía Nacional del Ecuador, y el medio digital El Ciudadano.

II. DESARROLLO DE LA PROPUESTA

A. Fuentes de Información

En este trabajo se ha usado como base el Diccionario Escolar Intercultural Bilingüe de kichwa-castellano - Ecuador [1]. Este diccionario es el producto de revisiones de los materiales producidos por instituciones de EBI, discusiones y aportes colectivos (DINEIB - ALKI) acerca de: nombres de los meses, nombres para los días de la semana, hora, minuto y segundo; de igual manera, se hallan los nombres de los planetas del sistema solar. Como se señala en [1], todo este trabajo se halla argumentado con evidencia histórica, lingüística, filosófica, cosmovisión, pedagógica, cultural, política y simbólica.

En lo relacionado al componente léxico del diccionario kichwa-castellano, consta de aproximadamente 1.920 entradas. Las entradas en el diccionario, constan casi todas en la segunda sección castellano kichwa. En esta parte se han precisado algunos significados, por ejemplo: hachuntinhachu, millón; huñu, diez mil; ayñi, arma; kaway, mirar; napaykuy, saludar; napay, estar en celo; taptana, ábaco andino; yupana, contador andino; tullpuy, teñir; llimpiy, pintar con colores; ñawi, ojo; unancha, bandera, emblema; uya, cara, rostro. No todo sinónimo tiene su lugar en la entrada respectiva, toda vez que no se trata de un diccionario de sinónimos kichwas.

En cuanto a la parte castellano-kichwa abarcar 2.800 entradas. Estas entradas no necesariamente se hallan todas en la parte primera (kichwa-castellano). Como muchos términos kichwas se hallan corrientemente ya en el habla castellana, así se registran en la entrada respectiva, ejemplo: inca, coya, quipo, carpa, cancha, canguil, chillca, chullpi, papa, quinoa... Otros vocablos se han kichuizado como: siwara, trigo, alvis, pircha, luway, kumun (cebada común, signo de pobreza de los kichwas), akichay, ransiw (<Francia), sawli (<sable). También se observan muchos términos del castellano antiguo que han sido adoptados como propios y la mayoría de hablantes (sin un conocimiento histórico de la lengua por supuesto) los asumen como kichwas: tayta, laya, harsia, amelga, milga, luway. Esta parte tiene número superior de entradas en relación a la parte kichwa-castellana. Por un lado, porque en varios casos, una entrada kichwa, dispone de tres o cuatro glosas castellanas (que demuestra el conocimiento más amplio del castellano), por ejemplo, sumak: lindo, bonito, hermoso, bello; churay: poner, colocar, ubicar, posicionar, etc. Todo lo contrario ocurre con el manejo de los sinónimos kichwas, con dificultad se maneja uno o dos sinónimos por término, en el caso de haberlo. Por otro lado, durante el registro de la parte castellana-kichwa, muchos términos eran necesarios consignarlos por lo menos en esta sección; además, el hecho práctico señala que el kichwa hablante va desde la glosa castellana a la del kichwa shimi. Los términos kichwas presentes en las variedades del Perú y Bolivia y otros que pertenecen al kichwa antiguo, procedentes del corpus neologismos ALKI, son aproximadamente 400 (de aproximadamente 3600 expresiones) [1].

Durante el proceso de recuperación y representación de conocimiento se detectarán los siguientes componentes de la microestructura léxica en [1]:

- entrada. Ésta se halla resaltada.
- etimología. Para la distinción y origen de ciertos términos (<*uspun > puzun).
- pronunciación. La pronunciación, entre corchetes [manga] <manka>.
- categoría gramatical (rumi, s.; puka, adj., wakin, det.).
- lugar de uso. Así: tu. Tungurahua, pi. Pichincha.

- glosa. Cada término tiene su traducción (rumi, piedra).
- descripción. La definición considera los términos que se hallan en la parte kichwa-castellano.
- ejemplificación. Cada entrada con su ejemplo respectivo y toda la estructura en cursiva y el término en cuestión resaltado.
- otras acepciones análogas. Éstas van numeradas con números índicos (1, 2, 3); su ordenamiento obedece a cuestiones lógicas e históricas que a su vez se tornan didácticos y se distribuyen en grupos de analogía; luego de este número el procedimiento es igual a la primera entrada.
- los sinónimos. Cierran la microestructura y con el número según corresponda a la acepción respectiva.
- términos de forma similar. Los formalmente iguales pero con significados distintos van en una entrada aparte (con números romanos: I makana, mazo; II makana, tipo de prenda; III makana, lugar para golpear).
- términos genéricos.- Se usan en la definición de ciertos nombres genéricos.

B. Tecnologías Utilizadas

Tienen relación directa en el desarrollo del proyecto las siguientes tecnologías:

- C. *Linked Data*. - Un enfoque que busca agilizar el tránsito de una red de documentos a una de conceptos vinculados, que permitan que las máquinas nos ayuden en la interpretación y procesamiento de información.
- D. *RDF*. - Es un método para expresar el conocimiento en un mundo descentralizado y es el fundamento de la web semántica, en el que las aplicaciones informáticas utilizan información estructurada distribuida por toda la web.[10]
- E. *OWL (Web Ontology Language)* extiende *RDF / RDFS* con vocabulario y semántica adicionales. Estos pueden verse como un marco para modelar (describir) todas las formas de datos a través del vocabulario y para permitir la interoperabilidad de los datos a través de conceptualizaciones compartidas y esquemas estandarizados y bien definidos [11].
- F. *Kichwa Spark Rest*: Está creado en el lenguaje de programación Java utilizando *Spark Framework*. Este es un framework web de desarrollo rápido inspirado en el framework *Sinatra* de Ruby y está basado en *Java 8 Lambda Expression Philosophy*. [7]
- G. *Knowledge Base*. - Una base de conocimiento es una base de datos utilizada para compartir y gestionar el conocimiento [12]. Se usara para recolectar, organizar y recuperar el conocimiento representado. El propósito no es necesariamente el almacenamiento de información sino más bien brindar alternativas de solución en función a los requerimientos planteados como alcance de este trabajo.
- H. *Knowledge Graph*. - Es una forma de conectar y unificar de manera significativa la información y hacerla interrogable de manera natural para las personas con el fin de lograr una web más inteligente [13].
- I. *Bot*.- Proviene de la palabra ‘robot’, es un software diseñado para cumplir una serie de tareas por su cuenta de manera independiente, es decir no necesita la ayuda de una persona. Este tipo de software funciona dentro de otras aplicaciones cumpliendo un rol específico.
- J. *ChatBot*. - Es un robot que permite la comunicación con el usuario, simulando una conversación con otra persona, con el propósito de responder a preguntas o brindar

información pertinente. Por tal razón los chatbots están presentes en aplicaciones de mensajería.

K. Protocolo seguido para estructurar contenido descrito en documentos no estructurados

1) Creación y captura

Yachakukkunapa Shimiyuk Kamu: Es un diccionario Kichwa – Castellano y Castellano – Kichwa. Es publicado por el Ministerio de Educación del Ecuador en el 2009, con el propósito de transmitir a la ciudadanía contenido relacionado a la historia y cultura de nuestros ancestros para de esta manera prevalecer esta lengua nativa. El formato de distribución es PDF.

Se llevaron a cabo tareas de extracción, organización y estructuración. Los conceptos recuperados de la sección Castellano – Kichwa se resumen a continuación:



La tabla 1 muestra un ejemplo de los contenidos obtenidos desde el diccionario castellano kichwa.

TABLE I. EJEMPLO - DICCIONARIO CASTELLANO KICHWA

<i>Castellano</i>	<i>Tipo</i>	<i>Kichwa</i>
Abono	Sustantivo	Wanu, isma
Casarse	Verbo	Sawarina
Grande	Adjetivo	Hatun, Kapak, Mama, Yaya

Desde la sección del Kichwa al Castellano se recuperan los siguientes conceptos:

Sustantivos	1053	Pronombres	14	Interjecciones	21
Verbos	612	Determinantes	16	Conjunciones	2
Adverbios	96	Preguntas	1	Expresiones	6
Adjetivos	197	Numerales	14	Onomatopeya	2
Ejemplos	4090	Sinonimos	1323	Comparativos	2

La tabla 2 muestra algunos ejemplos de traducción econtrado en el diccionario kichwa castellano.

TABLE II. DICCIONARIO KICHWA CASTELLANO

<i>Kichwa</i>	<i>Tipo</i>	<i>Castellano</i>	<i>Sinónimo</i>
Kutana	Verbo	Moler	Chamkana
Shuk	Número	Uno	-
Wayka	Determinante	Aglomeración	Tantariy
Amarak	Adverbio	Todavía	Manarak

Una vez identificado las fuentes de información es necesario convertir todos estos datos a un formato más manejable y accesible de tal manera que se los pueda manipular libremente para realizar proceso de limpieza, depuración y filtrado, razón por la cual estos datos pasaron a formar parte de documentos tipo excel.

2) *Transformación de contenido estructurado de un Grado de Conocimiento*

Transformar texto en lenguaje natural en una representación formal es uno de los objetivos importantes de la Web Semántica (SW). El concepto de una Web Semántica basada en datos vinculados (Linked Data) no es nuevo y, según Tim Berners-Lee, padre de la Web, se ha estado gestando durante mucho tiempo; el término se acuñó hace unos 20 años, y su propósito fue entretener una Web para que no solo vincule documentos entre sí, sino que también reconozca el significado de la información (datos y semántica) contenida en esos documentos.

La SW pretende facilitar la integración y recuperación de información heterogénea a escala global. La construcción de Grafos de Conocimiento, Knowledge Graphs (KGs), persigue esta idea, extraer entidades nombradas (cosas del mundo real) y sus relaciones desde información publicada en la Web. La Web Semántica y las tecnologías de LinkedData (LD) tienen el potencial de simplificar la integración y la interoperabilidad entre conocimiento heterogéneo, y proporcionan una forma natural de integrar diferentes formas de conocimiento. La intención de los KG es organizar la información con formatos procesables directamente por máquinas y enlazados semánticamente, de manera que los usuarios puedan navegar de una manera usable a través del KG creado. Se espera que de esta manera, las comunidades de Web Semántica, Procesamiento de Lenguaje Natural y Ciencia de Datos puedan complementarse y fortalecerse mutuamente en un ciclo de retroalimentación positiva.

El propósito de esta fase es diseñar un KG sobre entradas de lenguaje Kichwa y su significado en Castellano e Inglés. Se construyó una base de conocimiento usando el enfoque de datos enlazados. Las entidades, propiedades y relaciones que forman parte del KG producido están basadas en los procesos de extracción y estructuración descritos previamente. A partir de los elementos de información extraídos, se crearon triples RDF, potencialmente útiles para el KG resultante. El Grafo de Conocimiento creado se puede usar para muchas tareas diferentes. Como parte de este trabajo, se creará un prototipo de interfaz conversacional para la exploración del KG, con el objetivo de disponer de un vía de acceso de forma rápida a la información más importante al respecto de un término, así como diferentes rutas a través de las cuales podemos seguir explorando temas relacionados que, potencialmente, puede interesar al usuario.

3) *Enriquecimiento de entradas*

En este paso se enriquecen las entradas en kichwa a través del contenido del documento Kichwa-English-Spanish Dictionary. Este diccionario multilingüe está basado en el Kichwa utilizado en Ecuador, en su contenido lleva palabras que están escritas en el idioma Kichwa, Español e Inglés. A través de este diccionario, se enriqueció las entradas de [1] con significados en inglés; además a través de la expresión en inglés se anotó semánticamente el término con contenido obtenido desde WordNet. Se obtuvieron significados del concepto en inglés, así como las relaciones léxicas posibles de las palabras procesadas, para ello se utilizó WordNet.

Los resultados de este proceso son los siguientes:



La tabla 3 es un ejemplo de los contenidos que tiene el diccionario kichwa ingles español.

TABLE III. DICCIONARIO KICHWA-ENGLISH-SPANISH

<i>Kichwa</i>	<i>Ingles</i>	<i>Español</i>	<i>Tipo</i>
Yuyiy	Idea	Idea	Sustantivo
Zampa	Lazy	Peresoso	Adjetivo
Ñawpakpi	In front	Delante	Adverbio
Chayana	Get there	Llegar	Verbo

Los sustantivos, verbos, adjetivos y adverbios se agrupan en conjuntos de sinónimos cognitivos (synsets), cada uno expresando un concepto distinto [8]. Se utilizó el módulo NLTK que incluye WordNet con sus miles de palabras que están lógicamente relacionadas entre sí. A través del módulo desarrollado (escrito en Python) se obtuvo la información necesaria de cada una de las palabras Kichwa desde WordNet para posteriormente ser agregada a la base de datos.

4) Prototipado

Diferentes estudios sobre tendencias tecnológicas plantean que dentro de los próximos diez años, los *chatbots* y asistentes digitales como Siri y Alexa superarán a la página web como punto de contacto principal con los consumidores. Los *chatbots* buscan reemplazar aplicaciones móviles; y transformar el sector de atención al cliente a través de la transformación digital de las interfaces conversacionales. En este sentido, en el contexto de este trabajo, se diseñaron prototipos de explotación de la base de conocimiento creada a través del concepto de interfaces conversacionales. A continuación se describe de forma general los prototipos desarrollados.

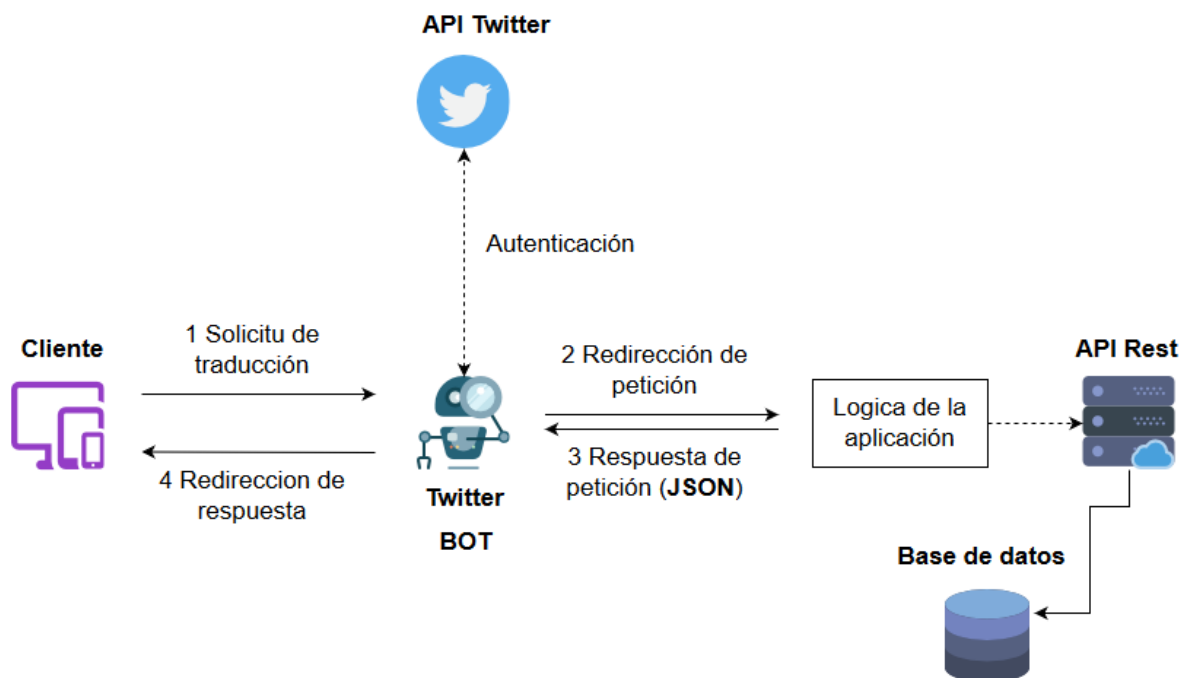


Fig. 2. Estructura del funcionamiento del bot Twitter.

a) Bot Twitter

- *Descripción:* La plataforma social Twitter nos permite compartir información de cualquier tipo y de una forma rápida a través de los mensajes cortos denominados tweets. Se desarrolló un bot, que tiene funcionalidad de responder a interacciones a través de tweets. Ver Figura 2. El tweet debe contener el contenido que necesita ser traducido, este será captado por un bot, se procesará, y el bot generará un mensaje de respuesta con el resultado del concepto en el idioma que se dese Kichwa/Castellano/Inglés.

- *Diseño:* La figura 2 muestra los componentes involucrados en el funcionamiento del bot en la plataforma Twitter.
- *Desarrollo:* Se utilizó Python como el lenguaje de programación, en su versión 3.4.

Para la autenticación con la Api de Twitter, se creó una aplicación en la plataforma Twitter. Esta aplicación nos permite obtener las credenciales necesarias para hacer uso de la Api: *Consumer Key*; *Consumer Secret*; *Access Token*; *Access Token Secret*. Para la explotación de la API, se usó la librería Tweepy que es compatible con el lenguaje Python. Se creó el bot encargado de realizar el proceso de traducción, este se mantiene en modo escucha de la publicación de nuevos tweets, siempre y cuando el tweet lleve uno de los hashtags preestablecidos, esto son:

- ✓ #kwUTPL.- Hashtag utilizado para la traducción de una palabra o frase kichwa al castellano.
- ✓ #casUTPL.- Hashtag utilizado para la traducción de una palabra o frase en castellano al kichwa.
- ✓ #engUTPL.- Hashtag utilizado para la traducción de una palabra o frase en kichwa al inglés.

Una vez que el bot detecta unos de estos hashtags, obtiene el texto del tuit y procede a consumir un servicio rest, que le brinda un resultado a partir de la palabra o frase que se encontró en el tuit. El resultado se lo obtiene en formato JSON, el mismo que es interpretado por el bot y a la vez se encarga de responder en un comentario al tuit que publico un determinado usuario.

- *Resultados preliminares:*
 - Si el usuario publica un tuit utilizando #kwUTPL, seguido de una frase o palabra kichwa, el bot responde con su respectiva traducción.
 - Si el usuario publica un tuit utilizando #casUTPL, seguido de una frase o palabra en castellano, el bot responde con su respectiva traducción al idioma Kichwa.

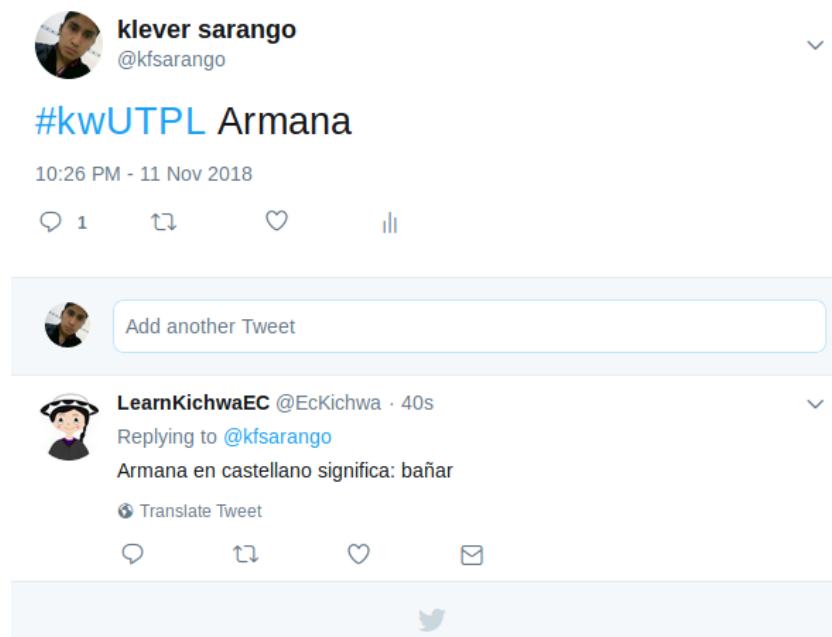


Fig. 3. Prototipo de Bot Twitter en funcionamiento.

La figura 3 muestra un ejemplo de respuesta que retorna el bot cuando detecta el hashtag #kwUTPL en la plataforma Twitter.

b) *Bot Telegram*

- *Descripción:* Telegram es una aplicación que ofrece el servicio de mensajería instantánea. Es una plataforma abierta que ofrece una Api que facilita a aplicaciones de terceros conectarse con la red. Esta Api permite la creación de Bots personalizables, lo cual facilita el proceso de creación de las diferentes funcionalidades que se le pueden asignar al bot. Como parte de este trabajo, se definió la creación de un segundo prototipo de interfaz conversacional. Se desarrolló un bot que permita la traducción de Kichwa a Castellano e Inglés (y viceversa) usando el grafo de conocimiento descrito anteriormente.

Se diseñaron los siguientes componentes:

- **Intentos y respuestas:** Los intentos son interpretados como una pregunta o entrada que el usuario realiza y las respuestas son acciones que el chatbot implementa en función del intento. Cuando iniciamos el bot por primera vez en Telegram, siempre el primer intento será el comando */start* como el punto de partida para ingresar a las funcionalidades que contiene el chatbot.
- **Entidades:** Los recursos de información representados en el grafo de conocimiento son las entidades desde las que se extrae la información requerida por el usuario.
- **Diálogos.** Se han diseñado flujos de interacción, con el objetivo de simular la comprensión de una conversación de manera similar a un diálogo en lenguaje natural.

- *Diseño:*

La figura 4 muestra cada uno de los componentes involucrados en el funcionamiento del bot en la aplicación de telegram.

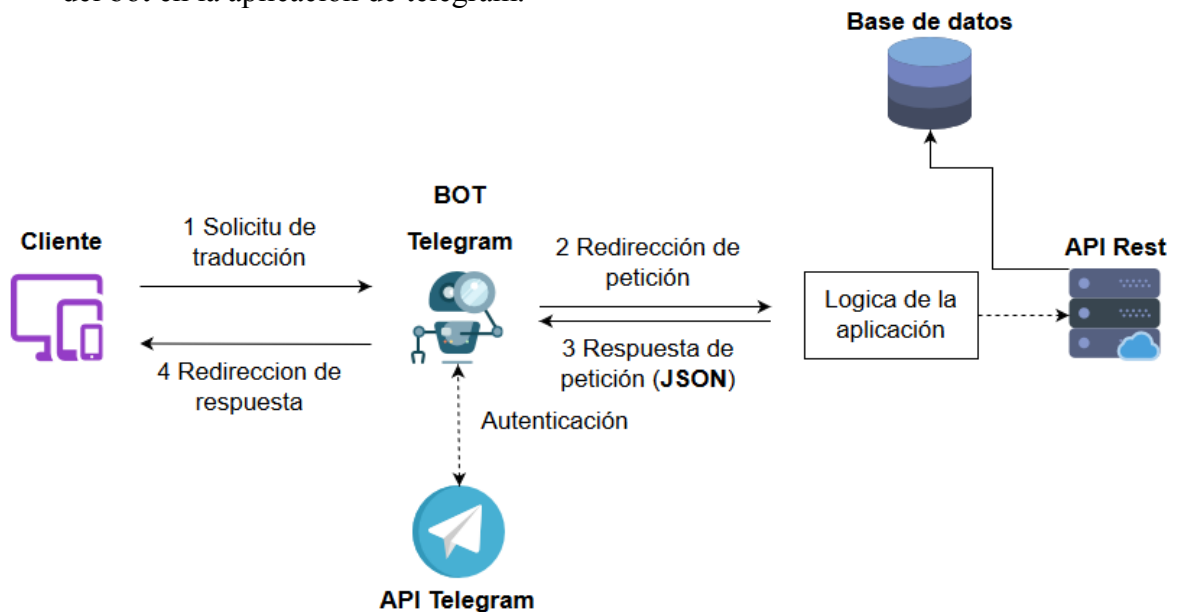


Fig. 4. Estructura del funcionamiento del bot Telegram.

- *Desarrollo:* El bot creado en este prototipo se denomina KichwaBot. Se utilizó Python como el lenguaje de programación, en su versión 3.4 respectivamente. Existen varias librerías para cada lenguaje de programación que permite utilizar la Api de Telegram, en nuestro caso se utilizó la librería Telepot que funciona en Python. Para la creación del menú de opciones se creó un menú de navegación que se visualiza al usuario. La función principal es la traducción, para la cual el usuario selecciona el lenguaje al que desea traducir y a continuación ingresa la palabra o frase a traducir.

El bot captura el contenido a traducir y procede a realizar la consulta a través del servicio rest. El resultado devuelto (en formato JSON) es procesado, se da formato y a continuación se comunica al usuario.

- *Resultados Preliminares:*
 - El usuario debe buscar y seleccionar el bot en la aplicación de Telegram para hacer uso de sus funcionalidades.
 - El usuario puede consultar la traducción de un término bien esta sea Kichwa; Castellano o Inglés.
 - La figura 4 Visualiza las diferentes funcionalidades que posee el bot.



Fig. 5. Bot desarrollado en Telegram.

III. CONCLUSIONES

La inclusión de nuevas tecnologías de la web permite administrar datos de manera eficiente en un contexto relacional con otros recursos disponibles en la web.

Se desarrolló este proyecto con el fin de incluir la lengua Kichwa en el mundo digital, para de esta manera compartir conocimiento y a la misma vez rescatar los valores culturales de los pueblos.

Existen diversas formas que se pueden implementar para promover la enseñanza de la lengua Kichwa por medio de la tecnología. Los usuarios de estas aplicaciones se sienten motivados en el aprendizaje de esta lengua, gracias a las nuevas metodologías que hacen de la enseñanza más divertida y entretenida.

Como trabajo futuro, se están prototipando nuevas características del chatbot relacionadas con el procesamiento de lenguaje natural, con la finalidad de poder asistir de mejor manera a los usuarios. Para el cual se define lo siguiente:

REFERENCIAS

- [1] SEPDI, "Kichwa Yachakukunapa Shimikuy Kamu," vol. 3, p. 276, Colección Runakay. Diccionario Escolar Intercultural Bilingüe de las Lenguas Ancestrales, No. 3. IICSAE 2016.
- [2] M. V. D. Mesias and A. G. W. David, "Estudio De Las Tecnologías De La Web Semántica Para El Desarrollo De Un Sistema De Traducción De Contenidos Español A Kichwa Para La Carrera De Sistemas Y Computación," p. 202, 2017.
- [3] M. Miguel Ángel Guzmán, "Aprendizaje Básico Del Idioma Kichwa Mediante Aplicación Informática Con Reconocimiento De Voz," p. 191, 2014.
- [4] Elizabeth Magaly Ayala Lliquín and B. E. U. Rodríguez, "Realidad Virtual Y Aumentada Como

- Herramienta De Aprendizaje Del Idioma Kichwa Para Niños,” p. 189, 2017.
- [5] L. E. M. Quinchuquí, “Aplicación Web De Enseñanza Multimedia Del Kichwa,” p. 174, 2014.
 - [6] A. T. Quezada, “Desarrollo e implementación de un Diccionario-Traductor del idioma Kichwa a Español para niños de educación básica de la escuela Jose Antonio Enrique de la ciudad de Loja,” pp. 1–38, 2015.
 - [7] Baeldung, “Building an API With the Spark Java Framework,” *March 1*, 2018. [Online]. Available: <http://www.baeldung.com/spark-framework-rest-api>.
 - [8] P. University, “WordNet A Lexical Database for English,” 2018. [Online]. Available: <https://wordnet.princeton.edu/>.
 - [9] Á. S. Lozada, «¿Qué es Linked Data?» [En línea]. Disponible en: http://www.semanticwebbuilder.org.mx/es/swb/Linked_Data.
 - [10] R. Mentor, «¿Qué es RDF? Introducción a RDF (Resource Description Framework)». [En línea]. Disponible en: <https://www.seofreelance.es/que-es-rdf-introduccion-a-rdf/>.
 - [11] N. Piedra, «LOD-CS2013 : Multilearning through a Semantic Representation of IEEE Computer Science Curricula», pp. 1945-1954, 2018.
 - [12] Techopedia, «Knowledge Base (klog)». [En línea]. Disponible en: <https://www.techopedia.com/definition/2511/knowledge-base-klog>.
 - [13] Gnos, «GNOSS Cognitive Platform: Extrae, conecta, analiza». [En línea]. Disponible en: <https://www.gnos.com/grafico-de-conocimiento>.