



# Desarrollo de un asistente virtual para brindar información y recolectar datos en las áreas de Salud, Educación y Economía como medida de mitigación al Covid-19

Lauro Enciso–Rodas<sup>1</sup>., Helard A. Becerra–Martínez<sup>2</sup>., Harley Vera–Olivera<sup>3</sup>., Rosa V. Encinas–Quille<sup>4</sup>., Ana R. Cárdenas–Maita<sup>4</sup>., Yanina Leon–Uscapi<sup>4</sup>., Meluni D. Palomino–Flores<sup>5</sup>., Jonathan R. Vasquez–Chicata<sup>6</sup>., Grover E. Castro–Guzmán<sup>4</sup>., Luis A. Quispe–Cartagena<sup>4</sup>., Jared León–Malpartida<sup>4</sup>. & Jhony L. Huallparimachi–García<sup>7</sup>.

1. Departamento de Informática, Universidad Nacional de San Antonio Abad del Cusco, Cusco, Perú.
2. University College Dublin, Irlanda.
3. Universidad de Brasilia - UnB, Brasilia, Brasil.
4. Universidad de São Paulo, São Paulo, Brasil.
5. Universidad politécnica de Puerto Rico, Puerto Rico, EEUU.
6. Universidad Andina del Cusco, Cusco, Perú.
7. Universidad de Puerto Rico - Recinto Mayaguez, Puerto Rico, EEUU.

Persona de contacto del proyecto: Lauro Enciso  
lauro.enciso@unsaac.edu.pe

## Resumen

El siguiente proyecto de investigación adaptativa presenta un asistente virtual (*chatbot*) llamado **Saminbot**, desarrollado con técnicas de procesamiento de lenguaje natural e inteligencia artificial. Con el objetivo de informar a la población y recolectar datos en temas relacionados a salud, educación y economía por medio de una conversación, dirigido a la población de Cusco. Esta conversación puede ser realizada mediante WhatsApp, Facebook Messenger y página web. Los datos recolectados estarán disponibles para realizar análisis y así dar soporte en la toma de decisiones a las autoridades del sector respectivo.

## Introducción

La pandemia del Covid-19 ha afectado la salud, educación y economía de las personas, alterando nuestras costumbres y estilos de vida. Nuestro país al igual que otros, ha hecho esfuerzos para cuidar la salud mediante el aislamiento social, apoyo económico a familias vulnerables y educación a distancia por TV. Sin embargo, estos dentro de estos sectores hay problemas que aún necesitan ser atendidos con urgencia, como es el seguimiento a las personas infectadas, detectar a personas económicamente vulnerables, e incluso hacer un monitoreo de los problemas que presentan los estudiantes que están llevando sus clases a distancia.

Diversas investigaciones vienen realizando esfuerzos para combatir la pandemia y sus efectos negativos. A pesar de los esfuerzos realizados una de las dificultades comunes a todos los investigadores ha sido la falta de acceso a datos reales para que sean procesados y luego analizados. Ya que estos datos aún son limitados o prohibidos para el acceso público. Por lo tanto es de vital importancia tener datos reales y confiables que representen el estado actual de la problemática de nuestro país en las áreas de salud, educación y economía, respetando la privacidad de las personas. Por otro lado, también la población Cusqueña necesita estar informada sobre temas relacionados a estas áreas ayudando a prevenir y/o apoyar su nuevo estilo de vida.

## Metodología experimental

La metodología es de rigor científico del tipo cuantitativo con alcance descriptivo. Para alcanzar los objetivos planteados siguen las siguientes etapas.

**Primera etapa:** Mapeamiento sistemático sobre chatbots, para la organización y consolidación de conceptos, modelos, herramientas disponibles, y posibilidades de aplicaciones de chatbots.

En esta etapa se realiza una investigación bibliográfica en 4 categorías principales: Aplicaciones para diversos sectores (salud, educación y economía), arquitecturas de chatbots, técnicas de Inteligencia Artificial (IA), y de procesamiento de lenguaje natural (PLN).

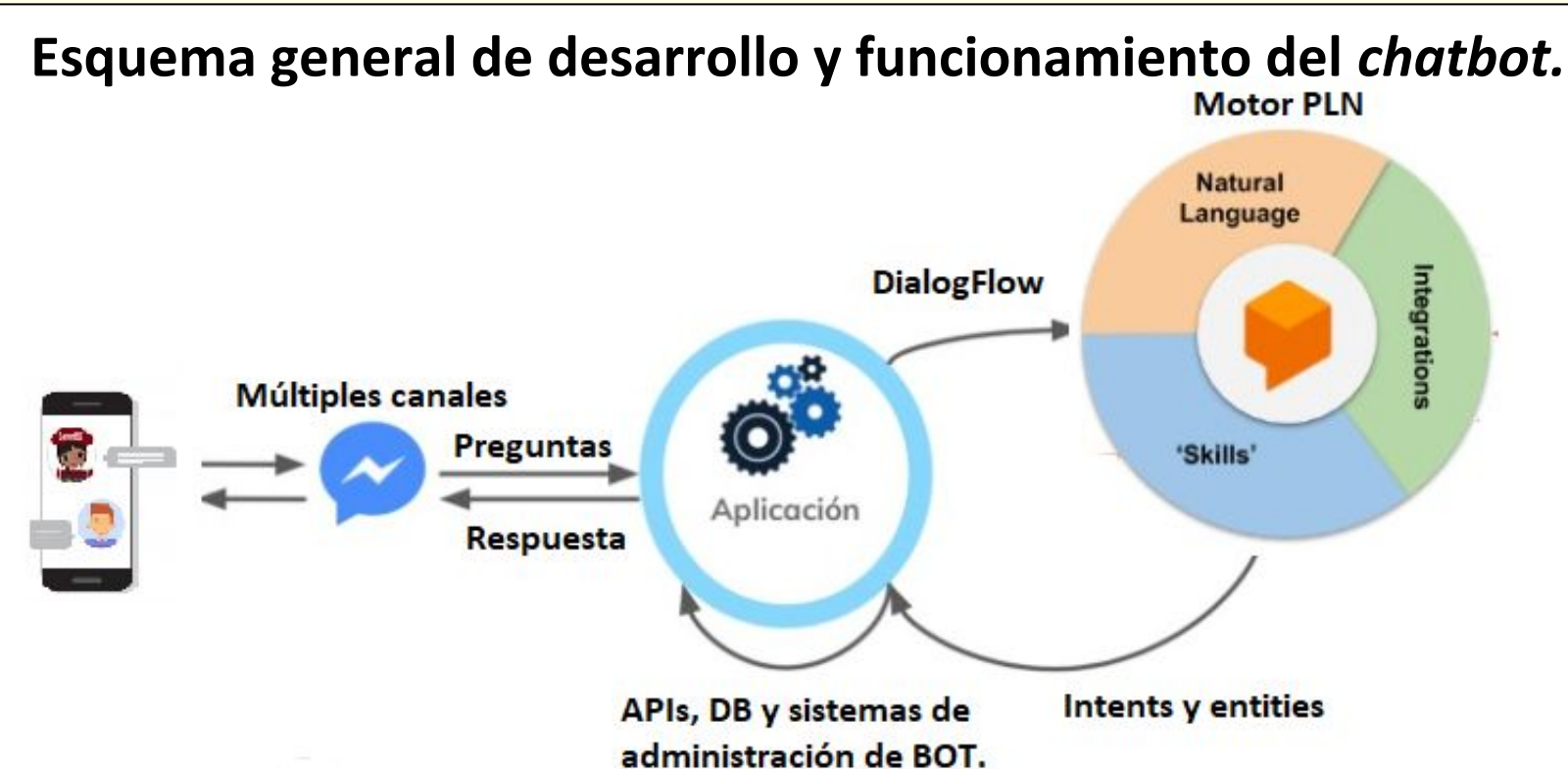
**Segunda etapa:** Se establecen los temas, estrategias y preguntas principales para ser evaluados e incluidos en el proyecto. Estas estrategias son específicas para cada área.

**Tercera etapa:** Definición de arquitectura de software del chatbot y datos a ser recolectados. La recolección de datos es realizada de acuerdo a la segunda etapa.

**Cuarta etapa:** Definición del modelo(s) a ser utilizado para el procesamiento del idioma español, utilizando las técnicas para la interacción con los usuarios.

**Quinta etapa:** Desarrollo del chatbot, el método utilizado para su desarrollo será por prototipado. Esto indica que el prototipo irá evolucionando de forma progresiva, a medida que todos los equipos vayan cooperando entre sí de forma paralela.

**Sexta etapa:** Resultados y análisis de resultados.



## Resultados esperados



Los resultados esperados son:

- Prototipo de asistente virtual (**Saminbot**) aplicado a las áreas de salud, educación y economía con funcionalidades propias para cada caso.
- Como resultado del uso del asistente virtual, se espera obtener tres *datasets* pre-procesados para el área de salud, educación y economía respectivamente para su posterior análisis. Los *dataset* serán de uso libre bajo la licencia *Creative Commons* Reconocimiento (CC BY).

Así mismo, gracias al uso del asistente virtual se espera obtener personas mejor informadas en las áreas de salud, educación y economía evitando información falsa (*fake news*) propagada por redes sociales principalmente. En el área de Educación se espera obtener los problemas presentados en esta modalidad de educación a distancia y brindarles información útil que les predisponga a aprender de forma dinámica con esta nueva opción tecnológica.



## Grupos de interés en el proyecto:

Los grupos de interés referente a los datos recolectados y a la información que brinda el asistente virtual en las áreas de salud, educación y economía, son:

- Interés en Informarse: Población Cusqueña en general; en los temas de salud, educación y economía.
- Interés en los datos: Comunidad científica;
- En el área de salud: Dirección regional de Salud (DIRESA), Red de Servicios de Salud Cusco.
- En el área de educación: La Unidad de Gestión Educativa Local Cusco (UGEL), Dirección Regional Educación Cusco (DRE CUSCO)
- En el área de economía: El Gobierno Regional.

La información recolectada será publicada de manera abierta para que las autoridades que se encuentren relacionadas con estas tres áreas puedan tomar decisiones en beneficio de la población cusqueña.

## Bibliografía

Abhishek S, Shivam S, Ramasubramanian K, (2019), Building an enterprise Chatbot: *Work with Protected Enterprise Data Using Open Source Frameworks*. Apress; 1st ed. Edición (13 Septiembre 2019). Comportamiento Adaptable de Chatbots Dependiente del Contexto - Juan Manuel Rodríguez , Hernán Merlino, Enrique Fernández.  
<http://revistas.unla.edu.ar/software/article/view/82>.