

Resultados del Análisis del procesamiento de lenguaje natural en la generación de agentes inteligentes conversacionales.

Results of analysis of natural language processing in the generation of intelligent conversational agents

Gil Santana Esparza.

Maestro en Comunicación Académica. Docente de la Academia de Ingeniería en Sistemas Computacionales. Instituto Tecnológico Superior de Pánuco (ITSP). México.

gil.santana@itspanuco.edu.mx

Gwendolyne Barrera Silva.

Licenciada en Educación. Docente de la Academia de Ingeniería en Gestión Empresarial. ITSP. México.

gwendybarrera@hotmail.com

Manuel Cornelio Torres Maza.

Maestro en Ciencias Administrativas. Docente de la Academia de Contador Público. ITSP. México.

manuel_torres_maza@hotmail.com

Resumen.

En este artículo se describen la propuesta y los resultados obtenidos derivados del proyecto de investigación denominado “Análisis del procesamiento de lenguaje natural con técnicas de Inteligencia Artificial en la generación de agentes inteligentes conversacionales”, el cual fue desarrollado en el Instituto Tecnológico Superior de Pánuco (ITSP), Veracruz, México. En este proyecto se planteó como objetivo investigar y analizar el procesamiento de lenguaje natural con técnicas de Inteligencia Artificial para diseñar agentes inteligentes conversacionales con funcionalidad de tutores virtuales que dieran como resultado final un prototipo software con enfoque de tutor para los alumnos del ITSP. Para lograr lo anterior se estudió y analizó el estado del arte del procesamiento del lenguaje natural, el estado del arte de los robots virtuales, las técnicas y paradigmas de programación de la Inteligencia Artificial orientados al diseño de robots virtuales. Se fortaleció con el análisis y estudio del lenguaje de programación AIML, para proponer un

Revista Avanzada Científica Enero – Abril Vol. 19 No. 1 Año 2016



Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported](https://creativecommons.org/licenses/by-nc-nd/3.0/).

diseño de la arquitectura de un robot virtual inteligente con capacidad de aprendizaje y el diseño del motor de inferencia y la respectiva base de conocimientos con enfoque de tutor virtual. Se complementó la investigación con los paradigmas de programación web PHP, XML y JavaScript. Se diseñó una metodología espiral evolutiva para el desarrollo del sistema software que contempla tres espirales con sus respectivos prototipos funcionales. Como resultado se obtuvo el prototipo que en las pruebas con el grupo experimental dio resultados satisfactorios y se propone para implantar en todos los grupos de tutoría individual del ITSP.

Palabras clave: Tutorías, índices de reprobación y deserción, lenguaje natural, Inteligencia Artificial, agentes conversacionales.

Abstract.

In this article the proposal and derived results of the research project "Analysis of natural language processing with Artificial Intelligence techniques in the generation of conversational intelligent agents for the area of Computer Systems Engineering" is described, which was developed Higher Technological Institute in Panuco (ITSP), Veracruz, Mexico. This project is aimed at investigating and analyzing the natural language processing with Artificial Intelligence techniques to design conversational intelligent agents functional virtual tutors to give final results in a software prototype of a virtual robot with a focus on tutor for students ITSP. To achieve this it is studied and analyzed the state of the art natural language processing, the state of the art virtual robots, techniques and programming paradigms of Artificial Intelligence oriented design virtual robots. It was strengthened by the analysis and study of the AIML language programming, to further propose an architecture design of an intelligent virtual robot capable of learning and inference engine design and the respective knowledge base approach virtual tutor. The research is complemented by the paradigms of web programming PHP, XML and JavaScript. An evolutionary spiral methodology for software development system that includes three laps with their respective functional prototypes designed. As a result the software prototype testing in the experimental group gave satisfactory results and intends to implement in all tutoring groups ITSP was obtained.

Keywords: Tutoring, failure and dropout rates, natural language, Artificial Intelligence, conversational agents.



Introducción.

El lineamiento para la operación del programa de tutorías de la Dirección General de Educación Superior Tecnológica (DGEST) (2011) menciona que su propósito es propiciar el mejoramiento de la calidad educativa contribuyendo a su formación integral. Por tal, la Tutoría es un proceso de acompañamiento grupal o individual que un tutor le brinda al estudiante durante su estancia en el ITSP con el propósito de contribuir a su formación integral e incidir en las metas institucionales relacionadas con la calidad educativa: elevar los índices de eficiencia terminal y bajar los índices de reprobación y deserción.

En base al análisis situacional realizado en la población estudiantil del ITSP, se detectó que el alumno en el proceso de aprendizaje y preparación de su carrera enfrenta dificultades emocionales y de decisión, en este caso las tutorías le han permitido esclarecer su rumbo y retomar a través de la asesoría sus proyectos profesionales. No obstante, se percibió en los resultados preliminares que los tutores no cubren las expectativas de tiempo y nivel de confianza de un número importante de estudiantes, lo cual puede repercutir en los índices de reprobación y en consecuencia la deserción a nivel Institucional. Lo anterior generó la necesidad de explorar nuevas formas de administrar la tutoría a estudiantes con necesidades personales o emocionales.

Para fortalecer lo anterior, el presente proyecto se planteó como objetivo analizar las teorías de procesamiento de lenguaje natural con técnicas de Inteligencia Artificial para la generación de agentes inteligentes conversacionales. El resultado del análisis de las teorías señaladas ha permitido diseñar un Tutor Virtual que se propone sirva de apoyo en el Programa Institucional de Tutorías desempeñando la función de atención psicológica a los alumnos canalizados desde la tutoría individual y grupal. De tal manera que el robot virtual desarrollará conversaciones entre un alumno tutorado y el software de emulación en lenguaje común y sobre temas de interés y necesidad del alumno atendido. Las conversaciones se llevan a cabo a través de tópicos desarrollados en función del curso de acción del intercambio de temas entre los actores.

Antecedentes.

Los sistemas computacionales diseñados para imitar el comportamiento humano requieren recursos de software especiales y no se pueden resolver eficientemente con los algoritmos tradicionales. En estos casos, se recurre a la Inteligencia Artificial que se inspira en lo natural para desarrollar técnicas que den soluciones a las tareas del ser humano. Una de esas tareas es la aplicación de sistemas

computacionales expertos que precisamente tienen como finalidad emular el comportamiento de un experto humano en una disciplina específica para solucionar un conjunto finito de problemas.

Actualmente las aplicaciones de sistemas que emulen la inteligencia humana con fines de aplicación en la actividad de tutoría a estudiantes universitarios todavía no están muy extendidas en México.

Existe un proyecto denominado A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) para la creación de robots virtuales que desarrollan conversaciones con los usuarios para propósitos generales. Algunos de estos robots virtuales, como CLAUDIO, han sido desarrollados para trabajar test de personalidad. También dentro de este proyecto, se encuentra DAVE que es un robot virtual que emula la función de un maestro de inglés para practicar el idioma.

En forma independiente al proyecto A.L.I.C.E. se encuentra ESL Mike que también es un robot virtual orientado a la práctica del idioma inglés en Internet. Así como también, en forma independiente existe EMOTE el cual es un robot virtual para fines académicos de la Universidad de Birmingham.

Por otro lado Mauricio Dorfman, docente investigador, se encuentra desarrollando el proyecto para un asistente virtual de clase en la educación universitaria en la Facultad de Ciencias Económicas de la Universidad de Buenos Aires.

En México, entre los antecedentes se encuentra la Universidad de Colima que tiene en desarrollo un Sistema Inteligente Conversacional para la Orientación Vocacional de los alumnos, el cual es un trabajo de grado de maestría en la Facultad de Telemática desarrollado por Ana Claudia Ruiz Tadeo. Este agente inteligente simula a un experto humano en orientación vocacional estableciendo una conversación con los usuarios para definir su vocación de acuerdo a sus capacidades e intereses personales.

En el Instituto Tecnológico de Cd. Madero Jesús Raymundo Domínguez ha presentado un proyecto de tesis para el diseño de un asistente virtual con diálogo emocional para proporcionar información sobre la Maestría en Ciencias en Ciencias de la Computación.

Como se puede apreciar en los párrafos anteriores, no existen proyectos documentados o divulgados que estén orientados a crear agentes conversacionales inteligentes con enfoque a robots tutores virtuales como apoyo al Programa Institucional de Tutorías, más bien están enfocados como agentes

conversacionales orientados a proporcionar información específica, a la práctica de idiomas, a desarrollar test de personalidad, a la orientación vocacional o con fines de entretenimiento.

Marco teórico.

Las raíces de la inteligencia artificial se hallan en la historia misma de la humanidad, y no sólo en su historia, también en los mitos y leyendas de muchas culturas. Si bien estas culturas no hablaban como tal de computadoras pensantes, sí se ha observado que se mencionan inteligencias artificiales o creadas por humanos. Es decir, que la idea de inteligencias creadas artificialmente siempre ha despertado interés en la raza humana. Se destaca en la historia de la inteligencia artificial el científico matemático Alan Turing que se interesaba en la posibilidad de crear una verdadera inteligencia artificial. Turing escribió un artículo en el cual trataba de contestar la pregunta sobre si las máquinas pueden pensar o no (Landa, 2007).

Actualmente no hay un consenso entre los científicos e ingenieros sobre una definición exacta y concisa que permita dirimir qué programas son o no inteligentes.

1. Inteligencia e Inteligencia Artificial.

Hofstadter (2008), menciona en su obra que una identidad inteligente, ya sea humana o computacional, debe responder a las situaciones de manera muy flexible y darle sentido a mensajes ambiguos o contradictorios. El aprendizaje automático es también condición necesaria para que un ente artificial pueda ser considerado inteligente. Si una máquina no es capaz de aprender cosas nuevas, difícilmente será capaz de adaptarse al medio, condición exigible a cualquier ser dotado de inteligencia (Landa, 2007).

2. Agente inteligente.

Se considera que una máquina es inteligente si piensa o se comporta racionalmente, basan sus técnicas en la lógica y en el concepto de agente. Según este enfoque, los actos de un agente inteligente deben basarse en el razonamiento y en las conclusiones obtenidas a partir de la información que se posee. Estos agentes tomarán la decisión más conveniente a la vista de esos datos y del tiempo del que disponen (García, 2013).

3. Robot virtual.

Revista Avanzada Científica Enero – Abril Vol. 19 No. 1 Año 2016



Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported](https://creativecommons.org/licenses/by-nc-nd/3.0/).

Tomando como base lo que comenta García (2013), un robot virtual es un software complejo que debe ser capaz de reconocer el lenguaje natural en el que hablan los humanos, y debe tener las siguientes capacidades para ser considerado en esta categoría: Reconocimiento de lenguaje natural, razonamiento, aprendizaje y representación del conocimiento

4. Procesamiento de lenguaje natural.

El procesamiento de lenguaje natural es una rama de la Inteligencia Artificial que se ocupa de las capacidades de comunicación de las computadoras con los humanos utilizando su propio lenguaje (García, 2013).

Sus aplicaciones son múltiples y variadas, como la corrección de textos, traducción automática, recuperación de la información, búsqueda de documentos o el reconocimiento y comprensión del lenguaje humano. El habla se asocia a una inteligencia superior, y para que una máquina sea capaz de reconocerla y también de construir frases tiene que ser capaz de realizar complejos análisis morfológicos, sintácticos, semánticos y contextuales de la información que recibe y de las frases que genera.

5. Sistemas basados en conocimiento.

Los sistemas basados en conocimiento se componen de dos partes principales: la base de conocimiento y el motor de inferencia. La base de conocimiento es la encargada de almacenar la información necesaria sobre el dominio de aplicación para poder ser capaz de llegar a conclusiones. El motor de inferencia es la parte que razona sobre la solución a un problema propuesto. A partir de una cuestión, el motor de inferencia busca información en la base de conocimiento y la relaciona hasta conseguir obtener una conclusión coherente con el problema que se le ha planteado (García, 2013).

6. A.I.M.L. (Artificial Intelligence Mark-up Language).

El lenguaje AIML, desarrollado por el Dr. Richard S. Wallace y la comunidad de software libre de The A.L.I.C.E. AI Foundation, posibilita la introducción de conocimiento dentro de los robots virtuales de manera que es el medio utilizado para la construcción de la inteligencia de los robots formando la base de conocimiento a través de un lenguaje de marcas derivado del lenguaje XML (Wallace, 2005).

El funcionamiento de AIML se basa en un modelo de impulso y de reacción a este impulso. La unidad básica en AIML se llama categoría. Una categoría es básicamente una pregunta o frase que un usuario escribe, lo que se conoce como patrón y una respuesta a la que se le conoce como plantilla. La pregunta es parte de la entrada y la respuesta es parte de la salida. Los patrones son los estímulos y las plantillas son las acciones.

7. Tutoría.

La Dirección General de Educación Superior Tecnológica (2011) (DGEST) establece que la Tutoría es un proceso de acompañamiento grupal o individual que un tutor le brinda al estudiante durante su estancia en el Instituto Tecnológico con el propósito de contribuir a su formación integral e incidir en las metas institucionales relacionadas con la calidad educativa; elevar los índices de eficiencia terminal, bajar los índices de reprobación y deserción. Dentro de las modalidades establece la tutoría individual la cual es una situación en la que se establece una relación directa cara a cara entre tutor y estudiante sobre cuestiones académicas individuales, de su situación personal, social o profesional; dándose una interacción, verbal y no verbal.

8. Tutor.

Se define al tutor, como el individuo que orienta, asesora y acompaña al estudiante durante su proceso de enseñanza-aprendizaje, con la perspectiva de una formación integral. Ayuda al tutorado a explorar sus capacidades, propugnando por la autoformación con base en el apoyo mutuo y en el trabajo en común (DGEST, 2011). De esta manera, el tutor se convierte en un acompañante durante un periodo educativo, que pretende guiar y apoyar en problemáticas educativas, sociales o psicológicas.

9. Tutorado.

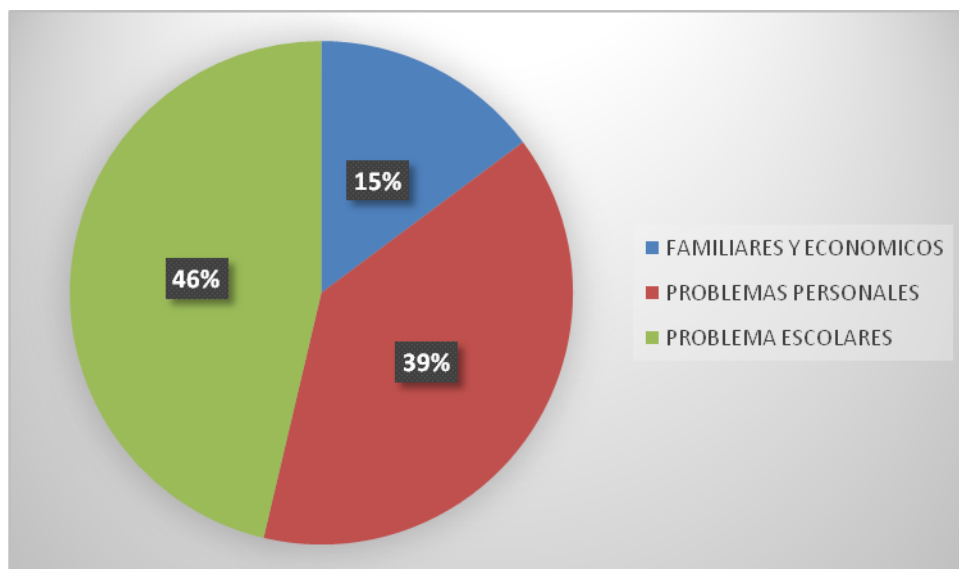
La DGEST (2011) menciona en su lineamiento que el tutorado es un estudiante que se responsabiliza de identificar sus necesidades académicas, administrativas y personales, respondiendo comprometidamente a la acción tutorial que le ofrece la institución. El estudiante tiene la disponibilidad e interés de mejorar su rendimiento, por ese motivo permite el acompañamiento, para desarrollar capacidades y prevenir los índices de reprobación o incluso de deserción.

Materiales y métodos.

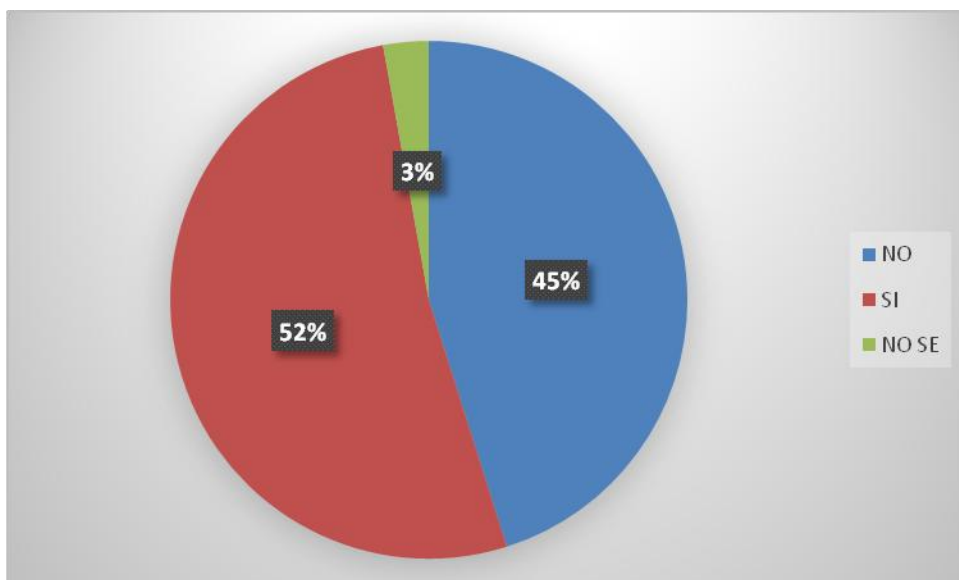
La metodología a seguir en este proyecto se basó en las teorías de Hernández, Fernández y Baptista (2014) para desarrollar una investigación documental combinada con una investigación de campo apuntando hacia la innovación tecnológica mediante la contribución de la documentación de las teorías investigadas y su aplicación en un prototipo software resultante de la investigación.

En la investigación documental se revisaron las teorías existentes sobre las técnicas para el procesamiento del lenguaje natural aplicando paradigmas de Inteligencia Artificial, así como también las técnicas para el diseño de robots conversacionales inteligentes (García, 2013).

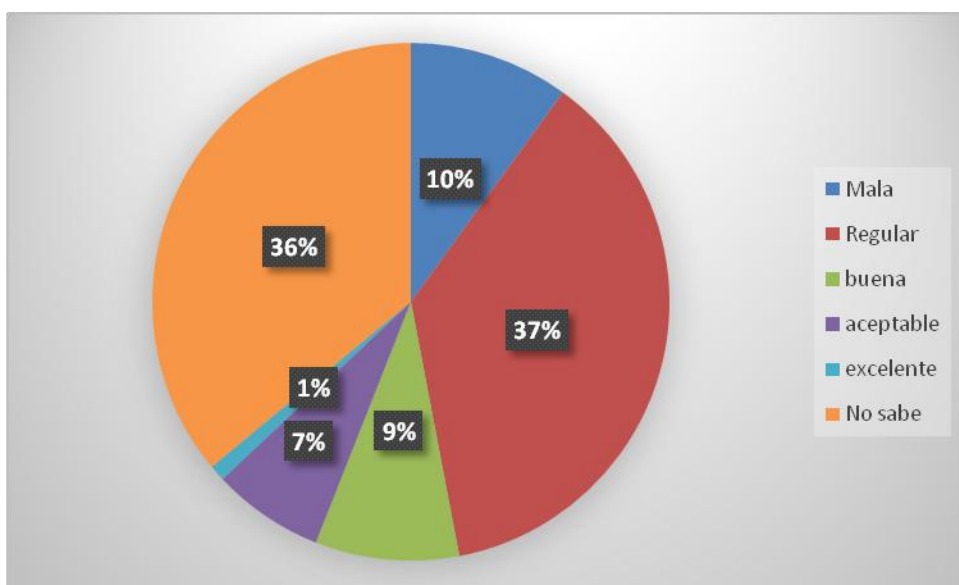
Para el enfoque final del prototipo software se recopiló información mediante entrevistas y cuestionarios a una muestra de 246 alumnos y 40 docentes del ITSP. Todo ello con la finalidad de obtener estadísticas que reflejen la situación actual y proyectar una alternativa de solución con aceptación en los alumnos.



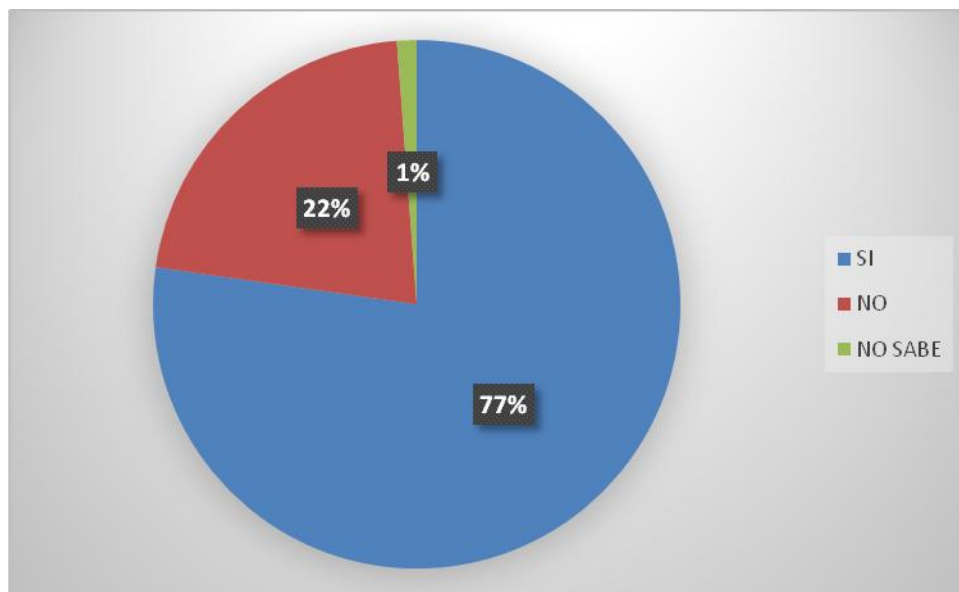
Grafica 1. Resultado de aplicar la encuesta a alumnos con la pregunta “¿Cuáles son los problemas más frecuentes por los cuales se requiere la tutoría?”



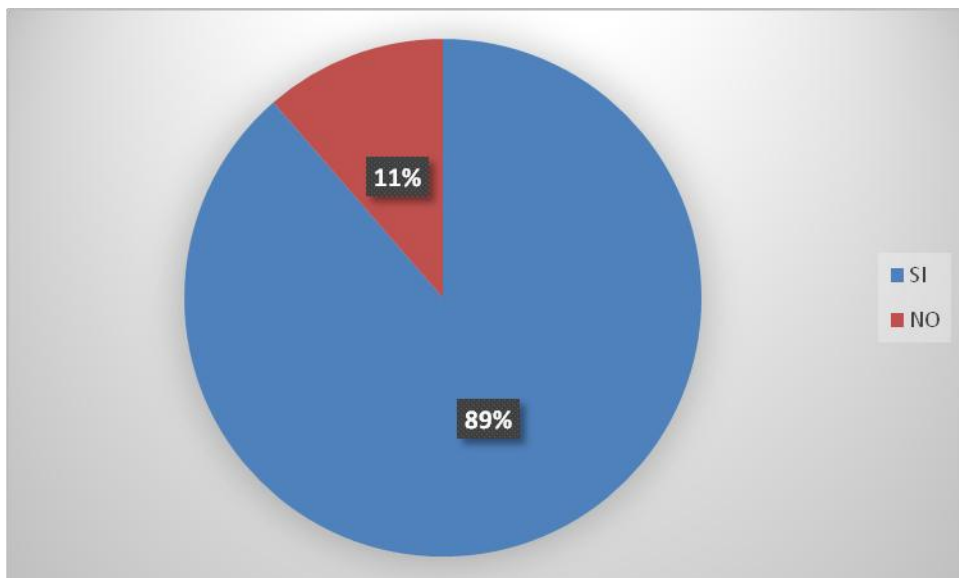
Grafica 2. Resultado de aplicar la encuesta a alumnos con la pregunta “¿Tienes la confianza para hablar de problemas personales con un docente tutor?”



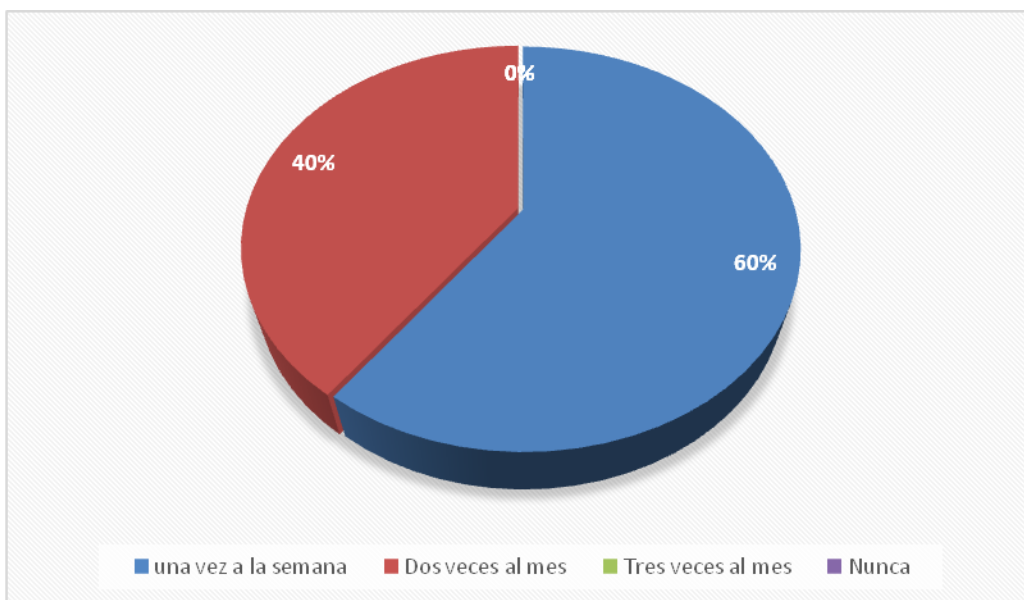
Grafica 3. Resultado de aplicar la encuesta a alumnos con la pregunta “¿Cómo consideras el procedimiento para solicitar una asesoría?”



Grafica 4. Resultado de aplicar la encuesta a alumnos con la pregunta “¿Confiarías en la implantación de un robot virtual como tutor especializado para los alumnos?”



Grafica 5. Resultado de aplicar la encuesta a alumnos con la pregunta “¿Estarías a favor de que la institución adquiriera un software que simule un tutor virtualmente?”



Grafica 6. Resultado de aplicar la encuesta a docentes con la pregunta “¿Con qué frecuencia los alumnos solicitan asesoría con un tutor?”

El análisis de las gráficas anteriores muestra que dentro de los problemas más frecuentes por los cuales se requiere la tutoría son del ámbito personal y escolar, pero un porcentaje importante de alumnos no tiene la confianza de tratar asuntos personales con un docente. Aunado a lo anterior, el procedimiento para solicitar una asesoría no es bien evaluada por los alumnos. Por otro lado, los docentes aceptan que los alumnos constantemente requieren asesoría y sin embargo en la institución sólo se cuenta con un tutor especializado dentro de la plantilla para atención de asuntos personales y emocionales.

Lo anterior generó la necesidad de explorar nuevas formas de administrar la tutoría a estudiantes con necesidades personales o psicológicas, para lo cual, se procedió a configurar el diseño básico del prototipo software del robot virtual conversacional con el enfoque de tutor virtual inteligente.

Para el desarrollo del software del robot tutor virtual se aplicó una metodología espiral evolutiva como lo plantean Pantaleo y Rinaudo (2014) que contempla las etapas de:

- **Análisis de requerimientos.**
Se estudiaron y analizaron detalladamente los requerimientos (necesidades) que se obtuvieron con las encuestas aplicadas en la institución detectando los requerimientos funcionales y no funcionales.
- **Diseño.**
Se establecieron las características para el servidor web, la base de datos y los módulos de programación del motor de inferencia y de la base de conocimiento, así como el módulo de reportes para análisis de información de la tutoría virtual.
- **Desarrollo.**
Se instaló y configuró el servidor web, se creó la base de datos y se programaron los módulos en los lenguajes AIML, PHP, XML y JavaScript.
- **Pruebas y ajustes.**
En esta etapa se aplicaron pruebas de funcionamiento y calidad de cada prototipo evolutivo.

La arquitectura de hardware y software (Pantaleo y Rinaudo, 2014) se basó en una arquitectura de tres capas cliente servidor. En la capa de datos se usó el paradigma de programación de Inteligencia Artificial AIML para diseñar la base de conocimientos. En la capa de lógica computacional se trabajó un servidor web Apache bajo los paradigmas de programación PHP combinado con XML interactuando con el paradigma AIML, esta capa es el enlace con la capa de presentación y la capa de datos. En la capa de presentación se usó XML con



JavaScript y código PHP para la interfaz de usuario incrustada en la plataforma virtual del Instituto.

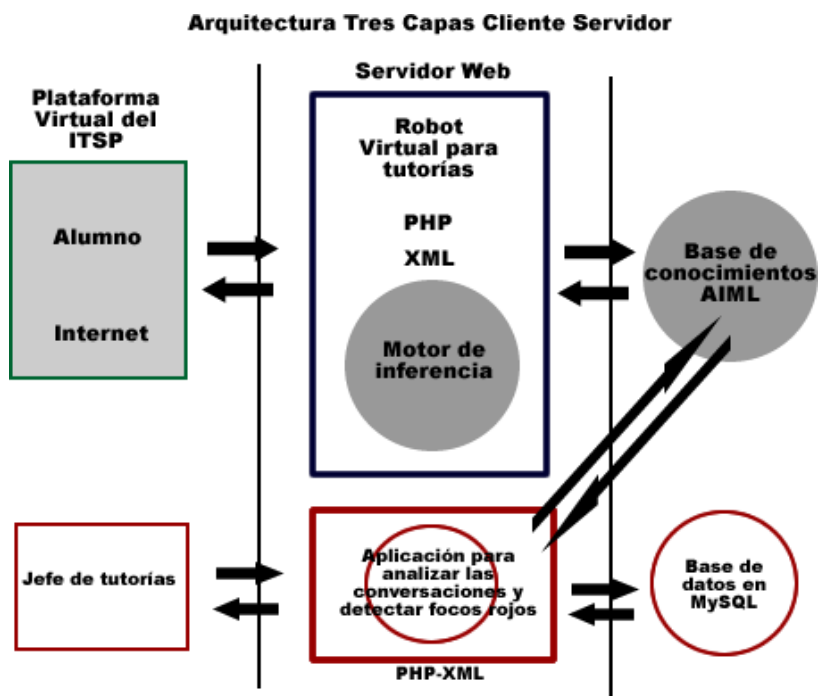


Figura 1. Arquitectura diseñada para el Robot Tutor Virtual del ITSP.

Resultado y discusión.

El resultado en esta etapa del proyecto es un prototipo software funcional que desarrolla conversaciones fluidas analizando las frases que el alumno introduce y llevando la conversación por el camino que normalmente haría un tutor especializado en asuntos personales y emocionales.



Figura 2. Pantalla de inicio de la conversación entre el alumno y el Robot Virtual.



Figura 3. Pantalla donde el alumno introduce el saludo inicial.



Figura 4. Pantalla de respuesta de saludo inicial.



Figura 5. Pantalla donde el alumno introduce el nombre con el cual quiere que se le identifique.



Figura 6. Pantalla donde el alumno introduce la respuesta a la carrera que estudia.

Internamente el software del Robot Virtual analiza cada frase tomando en consideración que el alumno puede contestar de muchas y muy variadas formas cada respuesta para lo cual realiza el análisis de las oraciones apoyándose en palabras clave que determinan la lógica y coherencia de la conversación. De tal manera que el alumno puede introducir el saludo con variantes como: Hola, ¿Hola cómo estás?, Hey!, Holaaaa!, etc. Y también puede introducir las distintas y múltiples formas de dar su nombre: Me dicen Roberto, Me llaman Roberto, Roberto, Soy Roberto, Pues mis amigos me dicen Roberto, Puedes llamarme Roberto, etc.

El Robot Virtual lleva el rumbo de la conversación y corrige el rumbo en caso de que el alumno salga del ámbito de la tutoría, que fue el conocimiento para el cual fue desarrollado.

Los beneficios obtenidos son:

- El aporte a la carrera de Ingeniería en Sistemas Computacionales de las teorías y técnicas de Inteligencia Artificial para la generación de agentes conversacionales.
- La contribución con una herramienta tecnológica como apoyo para el fortalecimiento del proceso de tutorías en los estudiantes.

Revista Avanzada Científica Enero – Abril Vol. 19 No. 1 Año 2016



Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported](https://creativecommons.org/licenses/by-nc-nd/3.0/).

- Fortalecimiento de la línea de investigación de inteligencia artificial.

Como trabajo a futuro se plantea como alternativa de seguir con el proceso de implantación del prototipo en ambientes reales a nivel institucional. Además se contempla para la siguiente etapa incluir la conversación a través de voz, tanto del alumno como del robot tutor virtual.

Conclusiones.

Se logró diseñar e implementar la arquitectura de un robot virtual incorporando una base de conocimiento relativa al área de tutorías conformando un prototipo de robot virtual conversacional que procesa el lenguaje natural a través de un motor de inferencia en PHP y el paradigma AIML. En esta primera etapa el agente conversacional se fortalece con un módulo externo en PHP y MySQL que graba las conversaciones en una base de datos convencional y que a través de reportes y gráficas permite enviar alertas para tomar acciones preventivas con los alumnos. Por lo anterior se puede afirmar que el diseño e implantación de agentes inteligentes conversacionales son un gran apoyo en las instituciones aprovechando que las generaciones actuales de jóvenes están inmersos en el mundo de la tecnología lo cual favorece el desarrollo de nuevos prototipos que involucren técnicas de Inteligencia Artificial.

Bibliografía.

- Cibelli, C. (2012). PHP Programación Web avanzada para profesionales. Primera edición. Buenos Aires, Argentina: Alfaomega.
- Dirección General de Educación Superior Tecnológica. (2011). *Lineamiento para la operación del programa de tutoría. Versión 1.0. Planes de estudio 2009-2010*. México, D.F., México: S.N.
- García, A. (2013). *Inteligencia Artificial. Fundamentos, práctica y aplicaciones*. Primera edición. México, D.F., México: Alfaomega.
- Hernández, R., Fernandez C., Baptista, P. (2014). *Metodología de la Investigación*. Sexta edición. México, D.F., México: McGrawHill.
- Hofstadter, D. (2008). *Yo soy un extraño bucle*. Primera edición. España: Tusquets Editores
- Landa, N. (2007). *Inteligencia Artificial*. Primera edición. Buenos Aires, Argentina: Gradi.
- Pantaleo, G. y Rinaudo, L. (2014). *Ingeniería de Software*. Primera Edición. Buenos Aires, Argentina: Alfaomega.
- SEP, DGEST. (2013). *Manual del tutor del SNIT*. México, D.F., México: S.N.

Wallace, R. (2005). *Be your own botmaster*. Segunda edición. Oakland, E.U.: ALICE A.I. Foundation, Inc.

Fecha de recepción: 25/01/2016

Fecha de aprobación: 18/04/2016

Revista Avanzada Científica Enero – Abril Vol. 19 No. 1 Año 2016



Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported](https://creativecommons.org/licenses/by-nc-nd/3.0/).