



# CHATBOT: CONCHITA COLEGIO LA CONCEPCIÓN IED

ELIYER JOHANNA CRUZ MATEUS

JHON ALEXIS TRIANA OLAYA

POLITÉCNICO GRANCOLOMBIANO

INGENIERÍA DE SISTEMAS

AUTÓMATAS, LENGUAJES Y GRAMÁTICAS

# INTRODUCCIÓN

Dentro del marco de la materia autómatas se plantea desarrollar como proyecto de aula, la aplicación de los autómatas en soluciones tecnológicas que faciliten la vida a los humanos. Por lo tanto nuestro trabajo consiste en implementar una solución tecnológica de comunicaciones en donde se pueda evidenciar la forma en que trabajan los autómatas. Para este fin se decidió desarrollar una Bot para la página web de un Colegio, de tal forma que esto de solución en parte a la comunicación y resolución de las dudas más frecuentes y mitigar los riesgos en la comunicación por medio de un chat automático, chatbot. El proyecto se implementará con una base de datos no SQL, el desarrollo del lenguaje se hará en el lenguaje de programación Python, la parte gráfica del chat se desarrollará con HTML, Javascript, CSS y la parte de los servicios con la API Angular.

# FORMULACIÓN DEL PROBLEMA

El colegio La concepción IED de la localidad de Bosa requiere un programa informático que resuelva las preguntas frecuentes de los padres de familia, estudiantes y personas interesadas en vincularse a la institución, sobre el Manual de convivencia, Sistema Institucional de Evaluación y la organización institucional; este programa deberá vincularse a la página web oficial del colegio.

# JUSTIFICACIÓN

El chatbot se desarrolla por la necesidad comunicativa de una población específica, en este caso la comunidad es un Colegio del Distrito donde sería muy útil tener una herramienta tecnológica que solucione las dudas que surgen en el día a día.

Esta solución tecnológica se enmarca en el desarrollo de Autómatas Finitos, de este modo se pueden usar la aplicación de la lógica matemática como una solución a un problema de comunicación. Teniendo en cuenta que la principal función de un autómata es “aceptar señales del medio que los rodea, cambiar de estado como consecuencia de estas señales y transmitir otras señales” \*.

\*Teoría de autómatas y Lenguajes formales. Autor Elena Jurado. [https://biblioteca.unex.es/tesis/Teorias\\_automatas.pdf](https://biblioteca.unex.es/tesis/Teorias_automatas.pdf)

# MARCO TEÓRICO

- Arquitectura Tecnológica de un Chatbot para la Gestión de la Información en una entidad superior

Autores: Ramírez Álamo, Yelitza Janeth, Carrizales Mamani, Jhon Rodrigo

Fuente: Universidad Peruana de Ciencias Aplicadas (UPC) ; Repositorio Académico - UPC

Año de publicación: 2020

# ARQUITECTURA TECNOLÓGICA DE UN CHATBOT PARA LA GESTIÓN DE LA INFORMACIÓN EN UNA ENTIDAD SUPERIOR

En este proyecto se propone un sistema cognitivo para mejorar la experiencia del usuario en la búsqueda de información académica con un chatbot. La diferencia entre los sistemas de búsqueda tradicionales y el sistema cognitivo propuesto es mejorar la experiencia del usuario (UX) a través de factores de optimización como el tiempo de respuesta, la facilidad de uso, la interfaz amigable y la interacción del usuario a través de los servicios cognitivos de computación en la nube. Esto puede ir más allá de la interacción entre un chatbot y un humano, ya que la experiencia del usuario es muy importante y puede definir el éxito o el fracaso de un sistema.



## MARCO TEÓRICO

- Retrospectiva al desarrollo de chatbots y procesamiento del lenguaje natural

Autores: Alemán Espinoza, José Wilfredo, Ruiz Vargas, Mario Rafael,  
Aguilar Juárez, Ernesto Alexander

Contribuidores: Universidad Francisco Gavidia (UFG)

Año de publicación: 2018



# RETROSPECTIVA AL DESARROLLO DE CHATBOTS Y PROCESAMIENTO DEL LENGUAJE NATURAL

Los chatbots se han vuelto cada vez más populares ; en parte gracias a la gran cantidad de herramientas disponibles para que los desarrolladores puedan implementar un agente conversacional en plataformas tan populares como Facebook ; en solo unos minutos. Sin embargo ; el desarrollo de entidades conversacionales no es nuevo ; y hay varios enfoques para proveer Procesamiento del Lenguaje Natural a dichas entidades ; de forma que se evite la trampa de implementar chatbots que no proveen ningún valor y evitar administrar un chatbot basado en una estructura if- then-else. El presente trabajo recopila trabajos históricos influyentes en el campo de la creación de entidades conversacionales ; o chatbots y habla sobre algunos enfoques para su implementación



The background of the slide is a dark blue-grey color. It features decorative white circuit-like lines with small circles at the ends, resembling a printed circuit board (PCB) layout. These lines are positioned along the left and right edges of the slide.

MARCO TEÓRICO

- Introducción Práctica Mogo DB

Autores: FaztWeb

URL: <https://www.faztweb.com/curso/mongodb>

Año de publicación: 2019



MARCO TEÓRICO

- Curso Python

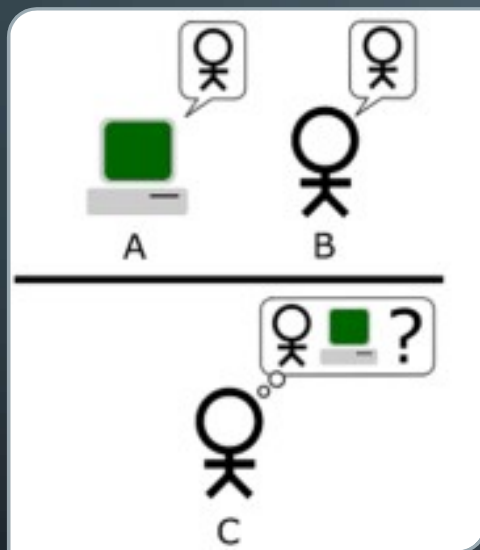
Autores: Píldoras informáticas

URL: <https://www.pildorasinformaticas.es/course/curso-python/>

Año de publicación: 2017



# MARCO CONCEPTUAL



## • Qué es un Chatbot<sup>[1]</sup>

Dentro de estos softwares de inteligencia artificial se encuentran los chatbots, que son bots especializados y creados para mantener conversaciones y ofrecer respuestas preconcebidas. Por lo tanto, un chatbot es un software que utiliza mensajes estructurados para emitir respuestas desde una máquina hacia un interlocutor humano.

Para que los chatbots sean eficientes, las empresas deben analizar sus procesos de atención al cliente y las conversaciones más comunes que se pueden desarrollar para construir las respuestas “enlatadas” y ofrecerlas en los momentos precisos a los usuarios/as.



## • Qué es MongoDB<sup>[2]</sup>

MongoDB es una base de datos de documentos que ofrece una gran escalabilidad y flexibilidad, y un modelo de consultas e indexación avanzado.

- MongoDB almacena datos en documentos flexibles similares a JSON, por lo que los campos pueden variar entre documentos y la estructura de datos puede cambiarse con el tiempo.
- Las consultas ad hoc, la indexación y la agregación en tiempo real ofrecen maneras potentes de acceder a los datos y analizarlos.



### • Qué es Python<sup>[3]</sup>

Python es un lenguaje de scripting independiente de plataforma y orientado a objetos, preparado para realizar cualquier tipo de programa, desde aplicaciones Windows a servidores de red o incluso, páginas web. Es un lenguaje interpretado, lo que significa que no se necesita compilar el código fuente para poder ejecutarlo, lo que ofrece ventajas como la rapidez de desarrollo e inconvenientes como una menor velocidad.

#### Ventajas:

- La cantidad de librerías que contiene, tipos de datos y funciones incorporadas en el propio lenguaje, que ayudan a realizar muchas tareas habituales sin necesidad de tener que programarlas desde cero.
- La sencillez y velocidad con la que se crean los programas.



ANGULAR

## • Qué es Angular<sup>[4]</sup>

Angular es un framework de desarrollo creado por Google para JavaScript. El lenguaje de la programación en Angular es TypeScript. La clave del éxito de Angular, es que permite crear aplicaciones web SPA (Single Page App) de forma sencilla.

### Ventajas:

- Facilidad a la hora de buscar documentación con una mayor coherencia en la forma de leer el código, así como un mantenimiento más sencillo de las aplicaciones web.
- Los componentes de la programación en Angular pueden reutilizarse en otros proyectos de la herramienta fácilmente, lo que hace que el desarrollo de las apps sea mucho más fluido. Los componentes son personalizados y funcionan perfectamente en navegadores modernos, y en bibliotecas o frameworks de JavaScript si trabajan con HTML.

## • Qué es Anaconda<sup>[5]</sup>

Es la Suite más completa para la Ciencia de datos con Python y que nos brinda una gran cantidad de funcionalidades que permitirán que desarrollemos aplicaciones de una manera más eficiente, rápida y sencilla.



Anaconda es una Suite de código abierto que abarca una serie de aplicaciones, librerías y conceptos diseñados para el desarrollo de la Ciencia de datos con Python. En líneas generales Anaconda Distribution es una distribución de Python que funciona como un gestor de entorno, un gestor de paquetes y que posee una colección de más de 720 paquetes de código abierto.



## Qué es Anaconda<sup>[5]</sup>

- Librería NumPy: es un paquete de Python que significa “Numerical Python”, es la librería principal para la informática científica, proporciona potentes estructuras de datos, implementando matrices y matrices multidimensionales. Estas estructuras de datos garantizan cálculos eficientes con matrices.
- Librería TFLearn: es una biblioteca de aprendizaje profundo modular y transparente construida sobre Tensorflow. Fue diseñado para proporcionar una API de alto nivel a TensorFlow con el fin de facilitar y acelerar las experimentaciones, sin dejar de ser totalmente transparente y compatible con ella.
- Librería nlkt: es el kit de herramientas de lenguaje natural es una de las bibliotecas NLP más conocidas y más utilizadas en el ecosistema de Python, útil para todo tipo de tareas, desde tokenización, hasta derivación, etiquetado de parte del habla.



# OBJETIVOS

## Objetivo General:

Implementar un Chatbot capaz de resolver preguntas frecuentes de la comunidad educativa del Colegio La Concepción IED (Padres de familia, estudiantes y personas interesadas en vincularse a la institución), sobre el Manual de convivencia, Sistema Institucional de Evaluación y la organización institucional.

## OBJETIVOS ESPECÍFICOS

- Crear una base de datos en Mongo DB del Manual de convivencia, , Sistema Institucional de Evaluación y la organización institucional para ser consultada a través de Python.
- Generar el código fuente del Chatbot en Python, para que responda las preguntas de la comunidad educativa con respuestas provenientes de la base de datos.
- Programar en Angular la aplicación web del Chatbot en donde ingresará la persona para realizar su(s) pregunta(s); esta aplicación web cumplirá con la posibilidad de vincularse a la página web del Colegio.

# DESARROLLO DEL AUTÓMATA

1. Para resolver las preguntas frecuentes de la comunidad educativa es necesario realizar una base de datos no relacional la cual contenga:

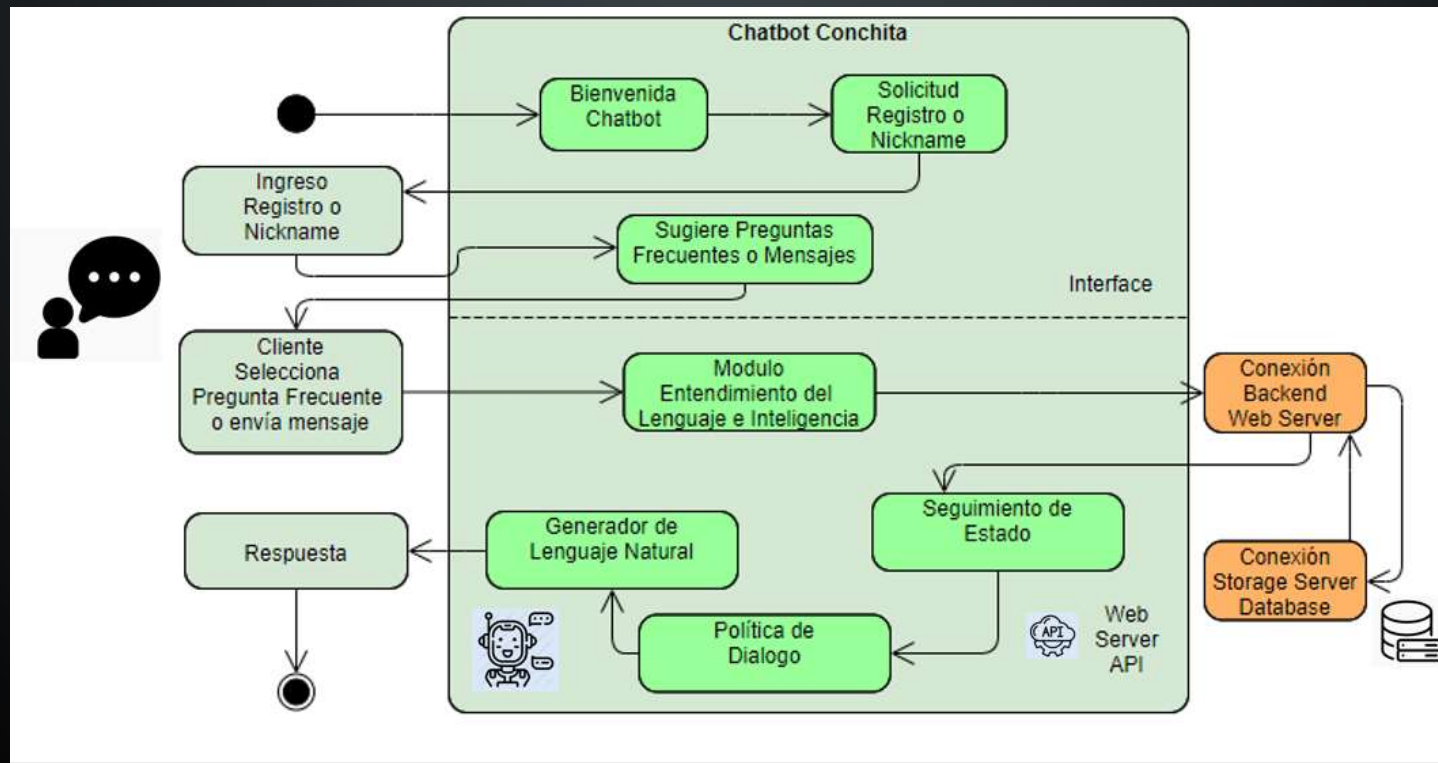
- Identificación y organización institucional (sedes y ubicación), horarios de atención de docentes, administrativos y equipo directivo.
- Manual de convivencia.
- Sistema Institucional de evaluación SIEE.

Para la implementación de esta base de datos se utilizará MongoDB.

## DESARROLLO DEL AUTÓMATA

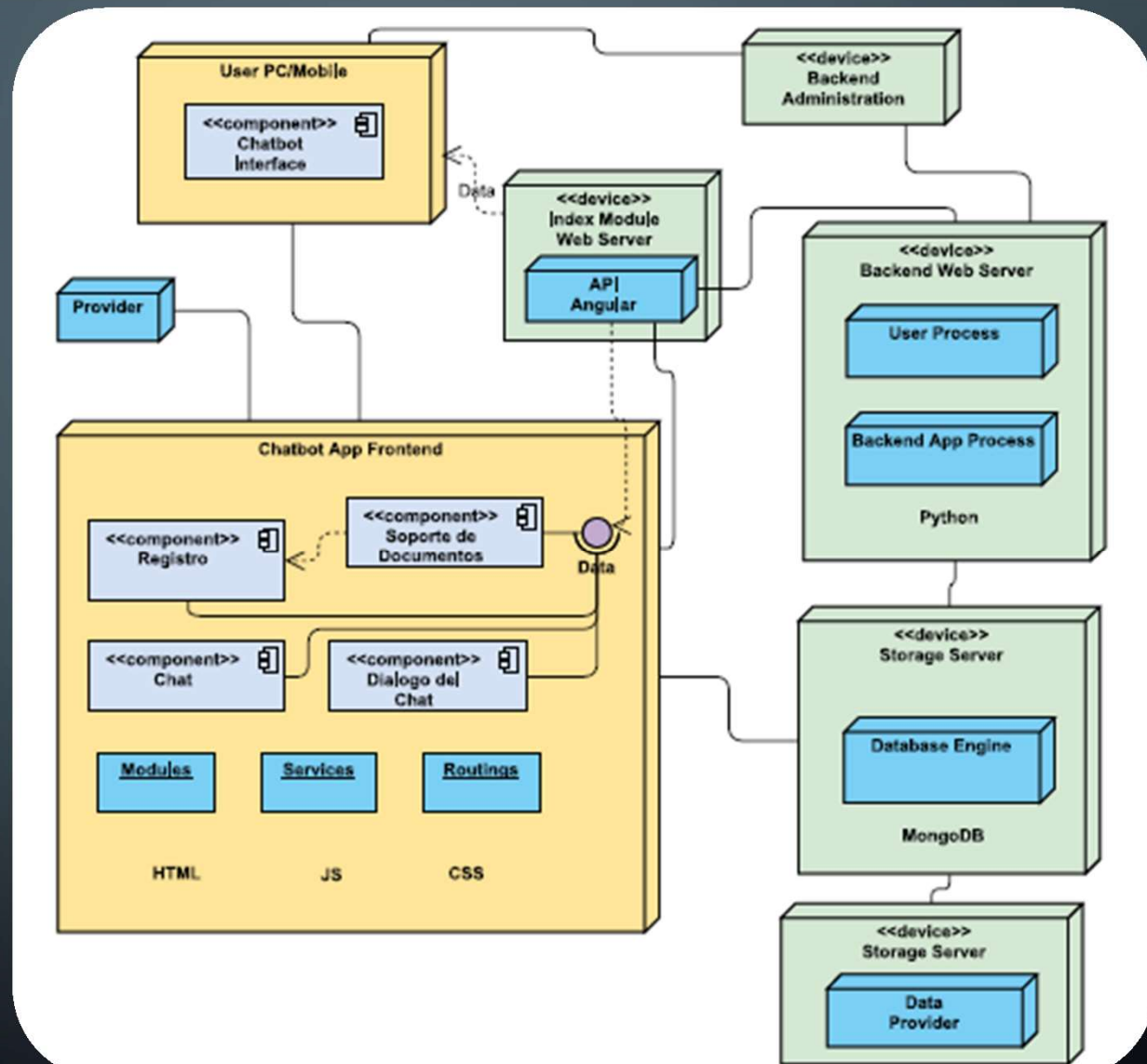
2. Para gestionar la base de datos es necesario realizar un programa utilizando el lenguaje de programación Python y se utilizarán las librerías PyMongo.
3. Para la creación del código fuente del Chatbot se utilizará el lenguaje de programación Python.
4. La API y el FrontEnd del Chatbot se realizará en Angular.

# DIAGRAMA DE ESTADOS DEL AUTÓMATA



# APLICACIÓN

<https://github.com/JhonTriana/chatbot>



# CÓDIGO FUENTE

```
Anaconda Powershell Prompt (anaconda3)

Collecting requests-oauthlib==0.7.0
  Downloading requests-oauthlib-1.3.0-py2.py3-none-any.whl (23 kB)
Collecting pyasn1==0.1.3
  Downloading pyasn1-0.4.8-py2.py3-none-any.whl (77 kB)
Collecting oauthlib==3.0.0
  Downloading oauthlib-3.1.0-py2.py3-none-any.whl (147 kB)
Building wheels for collected packages: termcolor, absl-py, wrapt
  Building wheel for termcolor (setup.py) ... done
  Created wheel for termcolor: filename=termcolor-1.1.0-py3-none-any.whl size=4835 sha256=7f77c8a58d07c21c6c84c9119a11942805dec8d1d72bb8b3078c1dd6c1eb0e
  Stored in directory: c:\users\mo\appdata\local\pip\cache\wheels\16\9c\5473df82468f958445479c59e784896fa24f4a5fc024b0f501
  Building wheel for absl-py (setup.py) ... done
  Created wheel for absl-py: filename=absl_py-0.9.0-py3-none-any.whl size=121935 sha256=7141a5e91353ac8261331bbd117b34837d206b6577aa80d4ca3c28cc0a9d2ae
  Stored in directory: c:\users\mo\appdata\local\pip\cache\wheels\1d\10\8e\2f79b924179ff6510933d63eb851bea01054fff262343b7a
  Building wheel for wrapt (setup.py) ... done
  Created wheel for wrapt: filename=wrapt-1.12.1-py3-none-any.whl size=19558 sha256=84193dc5d0b30f8bcfcd8473ca98618ec0b59b685e11c4fdaa7f3c1bbd9bbf
  Stored in directory: c:\users\mo\appdata\local\pip\cache\wheels\5f\fd\9e\b6cf5890494cb8ef0b5eaff72e5d55a70fb56316007d6dfe73
Successfully built termcolor absl-py wrapt
Installing collected packages: numpy, opt-einsum, six, google-pasta, termcolor, tensorflow-estimator, h5py, tensorboard-plugin-wit, grpcio, cachetools, pyasn1, rsa, pyasn1-modules, google-auth, chardet, idna, urllib3, requests, markdown, werkzeug, protobuf, absl-py, oauthlib, requests-oauthlib, google-auth-oauthlib, tensorboard, astunparse, gast, scipy, wrapt, keras-preprocessing, tensorflow
Successfully installed absl-py-0.9.0 astunparse-1.6.3 cachetools-4.1.0 chardet-3.0.4 gast-0.3.3 google-auth-1.14.0 google-auth-oauthlib-0.4.1 google-pasta-0.2.0 grpcio-1.28.1 h5py-2.10.0 idna-2.9 keras-preprocessing-1.1.0 markdown-3.2.1 numpy-1.18.2 oauthlib-3.1.0 opt-einsum-3.2.1 protobuf-3.11.3 pyasn1-0.4.8 pyasn1-modules-0.2.8 requests-2.23.0 requests-oauthlib-1.3.0 rsa-4.0 scipy-1.4.1 six-1.14.0 tensorboard-2.2.1 tensorboard-plugin-wit-1.5.0 tensorflow-2.2.0rc3 tensorflow-estimator-2.2.0 termcolor-1.1.0 urllib3-1.25.9 werkzeug-1.0.1 wrapt-1.12.1
(chat) PS C:\Users\MO>
(chat) PS C:\Users\MO> pip install numpy
Requirement already satisfied: numpy in c:\users\mo\anaconda3\envs\chat\lib\site-packages (1.18.2)
(chat) PS C:\Users\MO> pip install tflearn
Collecting tflearn
  Downloading tflearn-0.3.2.tar.gz (98 kB)
  Requirement already satisfied: numpy in c:\users\mo\anaconda3\envs\chat\lib\site-packages (from tflearn) (1.18.2)
  Requirement already satisfied: six in c:\users\mo\anaconda3\envs\chat\lib\site-packages (from tflearn) (1.14.0)
Collecting Pillow
  Downloading Pillow-7.1.1-cp38-cp38-win_amd64.whl (2.0 MB)
  Building wheels for collected packages: tflearn
  Building wheel for tflearn (setup.py) ... done
  Created wheel for tflearn: filename=tflearn-0.3.2-py3-none-any.whl size=128212 sha256=8ddde648b0373ceafcd8c6ecf780c6579b2134f040152db79643bb68a5dc38a7
  Stored in directory: c:\users\mo\appdata\local\pip\cache\wheels\67\df\9e\ec81a96bf237a5ffbeffb519138efebb3fe651be57be2d1f4
Successfully built tflearn
Installing collected packages: Pillow, tflearn

C:\Users\MO\Desktop\POLI\Autómatas\Proyecto\Chatbot python - Notepad++
Archivo  Editar  Buscar  Vista  Codificación  Lenguaje  Configuración  Herramientas  Macro  Ejecutar  Plugins
Ventana  ?
MinSql.py  Objetos Json BD_CLC  Chatbot python
1  modelo = tflearn.DNN(red)
2  modelo.fit(entrenamiento,salida,n_epoch=1000,batch_size=10,show_metric=True)
3  modelo.save("modelo.tflearn")
4
5  def naubBot():
6      While True: #Ciclo para mantener activo el chat durante la interaccion con el usuar:
7          entrada= input("Tu: ")#Creacion de la variable que guarda lo que el usuario dig:
8          palabrasIngresadas = [0 for _ in range(len(palabras))] #Creacion de un array qui
9          entradaProcesada = nltk.word_tokenize(entrada) #Creacion de la variable que guar
10         entradaProcesada = [stemmer.stem(palabra.lower()) for palabra in entradaProcesad
11
12
length: 881 lines: 12 Ln: 12 Col: 21 Sel: 0|0 Windows (CR LF) UTF-8 INS
```



# BASE DE DATOS EN MONGO DB

Colecciones	Documentos
horizonteInstitucional	Misión Visión Fundamentos Institucionales Objetivo general
identificacionInstitucional	Colegio La Concepción IED
organizacionAdministrativa	Sedes Rectoría Secretaría Coordinación Orientación Profesional de apoyo NEE (Necesidades educativas especiales) Horarios de los estudiantes Jornadas, docentes y horarios Personal apoyo administrativo
perfilesInstitucionales	Perfil del estudiante concepcionista Perfil del docente y directivo docente concepcionista Perfil del padre de familia y/o acudiente concepcionista
simbolosInstitucionales	PEI LEMA Nuestro himno



```

C:\Windows Server 2008 R2
2020-04-17T15:35:28.831-0500 I CONTROL [initandlisten] db version v4.2.5
2020-04-17T15:35:28.831-0500 I CONTROL [initandlisten] git version: 2261279b51ea13df08ae708ff278f0679c59dc32
2020-04-17T15:35:28.831-0500 I CONTROL [initandlisten] allocator: tcmalloc
2020-04-17T15:35:28.832-0500 I CONTROL [initandlisten] modules: none
2020-04-17T15:35:28.832-0500 I CONTROL [initandlisten] build environment:
2020-04-17T15:35:28.832-0500 I CONTROL [initandlisten] distmod: 2012plus
2020-04-17T15:35:28.832-0500 I CONTROL [initandlisten] distarch: x86_64
2020-04-17T15:35:28.833-0500 I CONTROL [initandlisten] target_arch: x86_64
2020-04-17T15:35:28.833-0500 I CONTROL [initandlisten] options: {}
2020-04-17T15:35:28.859-0500 I STORAGE [initandlisten] Detected data files in C:\data\db\ created by the 'wiredTiger' storage engine, so setting the active storage engine to 'wiredTiger'.
2020-04-17T15:35:28.859-0500 I STORAGE [initandlisten] wiredtiger_open config: create,cache_size=5340M,cache_overflow=(file_max=0M),session_max=33000,eviction=(threads_min=4,threads_max=4),config_base=false,statistics=(fast),log=(enabled=true,archive=true,path=journal,compressor=snappy),file_manager=(close_idle_time=100000,close_scan_interval=10,close_handle_minimum=250),statistics_log=(wait=0),verbose=[recovery_progress,checkpoint_progress],
2020-04-17T15:35:29.057-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:56838][14272:140717700046416], txn-recover: Recovering log 3 through 4
2020-04-17T15:35:29.164-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:163753][14272:140717700046416], txn-recover: Recovering log 4 through 4
2020-04-17T15:35:29.246-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:246692][14272:140717700046416], txn-recover: Main recovery loop: starting at 3/174464 to 4/256
2020-04-17T15:35:29.372-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:371818][14272:140717700046416], txn-recover: Recovering log 3 through 4
2020-04-17T15:35:29.484-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:483798][14272:140717700046416], txn-recover: Recovering log 4 through 4
2020-04-17T15:35:29.550-0500 I STORAGE [initandlisten] WiredTiger message [1587155729:549759][14272:140717700046416], txn-recover: Set global recovery timestamp: (0, 0)
2020-04-17T15:35:29.831-0500 I RECOVERY [initandlisten] WiredTiger recoveryTimestamp: Ts: Timestamp(0, 0)
2020-04-17T15:35:29.898-0500 I STORAGE [initandlisten] Timestamp monitor sta

```

```

C:\Windows Server 2008 R2
2020-04-17T15:35:29.937-0500 I CONTROL [initandlisten] ** addresses it should serve responses from, or with --bind_ip_all to
2020-04-17T15:35:29.942-0500 I CONTROL [initandlisten] ** bind to all interfaces. If this behavior is desired, start the
2020-04-17T15:35:29.943-0500 I CONTROL [initandlisten] ** server with --bind_ip 127.0.0.1 to disable this warning.
2020-04-17T15:35:29.945-0500 I CONTROL [initandlisten]
---
Enable MongoDB's free cloud-based monitoring service, which will then receive and display metrics about your deployment (disk utilization, CPU, operation statistics, etc).

The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product improvements and to suggest MongoDB products and deployment options to you.

To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
---
> use CLC
switched to db CLC
> show collections
asignacionHorarios
horizonteInstitucional
identificacionInstitucional
organizacionAdministrativa
perfilesInstitucionales
simbolosInstitucionales
> db.identificacionInstitucional.find().pretty()
{
  "_id" : ObjectId("5e988033723f2cd978c2dc14"),
  "nombre" : "Colegio La Concepción IED",
  "DANE" : "11100186758",
  "NIT" : "8300490920",
  "resolución" : [
    "Básica Primaria 7453 de 13/10/1998",
    "Básica secundaria 070215 de 4/01/2010",
    "Media Vocacional 07-123 de 25/09/2017"
  ]
}

```

# WEBGRAFÍA

[1] Jesús Charlán. Qué es un chatbot y para qué sirve. Economía digital

4 agosto, 2018. Recuperado de:

<https://www.esic.edu/rethink/2018/08/04/que-es-un-chatbot-y-para-que-sirve/>

[2] ¿Qué es MongoDB? Recuperado de: <https://www.mongodb.com/es/what-is-mongodb>

## WEBGRAFÍA

[3] Miguel Ángel Álvarez. Qué es Python. Recuperado de:  
<https://desarrolloweb.com/articulos/1325.php>

[4] ¿Qué es y para qué se utiliza la programación en Angular?. AZPE Informática. Recuperado de: <http://azpe.es/blog/que-es-programacion-en-angular/>

[5] ¿Qué es Anaconda Distribution?. Desde Linux. Recuperado de:  
<https://blog.desdelinux.net/ciencia-de-datos-con-python/>