

**Arquitectura del sistema**

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre del proyecto: | Plataforma de prevención y monitorización de discriminación del estado de Veracruz. | | |
| Alias del Proyecto: | PREDEV | | |
| Versión | 2.1 | Fecha última modificación: | 16/03/2019 |

El documento muestra la arquitectura de la plataforma PREDEV en la cual se muestra la estructura general y funcionalidad de este sistema. Se muestra a detalle la acción de cada uno de los módulos y las tecnologías que se requieren para realizar cada una de las acciones.

La plataforma PREDEV esta compuesta por cinco módulos los cuales realizarán funciones de extracción, análisis, guardado y mostrar la información obtenida de la web en general y de redes sociales como Facebook enfocando la extracción en la paginas oficiales de periódico o institución, los módulos son Extracción de datos, Backend , análisis de datos, base de datos y Frontend.

La plataforma estará automatizada para que cada uno de los módulos realice sus funciones y muestren la información en estadísticas en tiempo real para los usuarios o instituciones .

Para realizar cada uno de los módulos, se toma en cuenta la tecnología necesaria para dichas funciones, entre las cuales podemos mencionar lenguajes de programación como PHP , Python , APIs de análisis de texto y framework para la parte visual de la plataforma.

A continuación, se muestra la descripción de cada uno de los módulos de la plataforma PREDEV y la forma grafica para entender la funcionalidad del sistema.

# 

# **La arquitectura del sistema PREDEV se compone de las siguientes módulos:**

* Extracción de datos: esta acción será la mas importante, en la cual se extraerán los datos de las paginas web enfocadas a la discriminación en la zona del estado de Veracruz, así como de la red social Facebook en la cual también se extraerá la información de paginas oficiales de noticias.
* Backend: en esta acción se realizara la parte mas importante de la extracción en donde se trabajara con lenguajes como PHP o Python según se requiera. Aquí extraerá las noticias mas importantes y recientes que sean relacionadas con la discriminación. Se estructura la información para posteriormente ser analizada y determinar las estadísticas de cada zona del estado de Veracruz.
* Analizar los datos: una vez que la información se tenga extraída de las paginas mas importante y estructurada, se analizara y se guardaran los porcentajes tomando en cuenta la noticia y en donde ocurrió la noticia para así aumentar porcentaje por zona. Se utilizaran APIs para realizar esta acción ya que estas herramientas nos ofrecen diccionarios de palabras y funciones para analizar texto y obtener resultados con base a ello.
* Base de datos: cuando la información ya este extraída y analizada en este modulo se guardara en la base de datos MySQL de la plataforma que estará ubicada en un servidor externo, al ser almacenamiento masivo de información se guardara en la nube para tener la plataforma siempre optimizada y sin problemas de retardo o perdida de información.
* Frontend: este modulo será muy importante, a que aquí se mostrara toda la información obtenida anteriormente pero se mostrara con porcentaje mediante un mapa del estado de Veracruz utilizando lenguajes de lado del cliente y frameworks como Boostrap para la parte visual y el cual se alimentara de las noticias en tiempo real para mostrarse y que genere estadísticas certeras.