



Acceso Inteligente a la Información

Sistema de búsqueda y recomendación de noticias

- Daniel Moreno Soto
- Francisco Jesús Belmonte Pintre



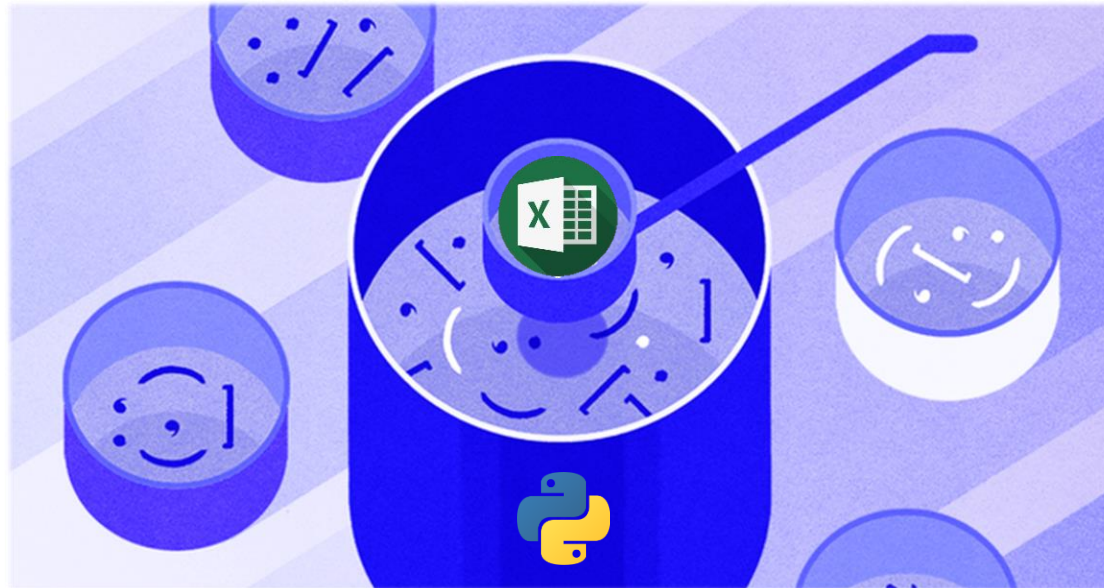
Herramientas utilizadas



Herramientas utilizadas

Python + BeautifulSoup

- Web Scraping
- Generar un archivo “csv” con toda la información de las noticias.



Herramientas utilizadas

Python + BeautifulSoup



Sección o
Categoría

Título

Autor

Descripción

Enlace
original

Contenido

	A	B	C	D	E	F
1	secc	titulo	autor	descr	enlaceTema	contenido
2						
3	España	La primera reunión de Ayuso y Mon	OKDIARIO	Isabel Díaz Ayuso y Roc	https://okdiario.com/espana,	['Isabel Díaz Ayuso y Rocío Monasterio han tenido este jueve
4						
5	España	Los separatistas convocan manife	OKDIARIO	Los separatistas ponen	https://okdiario.com/espana,	['Los separatistas ponen las instituciones públicas catalanas :
6						
7	España	Sánchez vuelve a excluir a Abasca	Joan Guirado	El presidente del Gobie	https://okdiario.com/espana,	['El presidente del Gobierno en funciones, Pedro Sánchez, vo
8						
9	España	Casado no tiene «ningún impedim	OKDIARIO	Pablo Casado, líder del	https://okdiario.com/espana,	['Pablo Casado, líder del PP, ha afirmado que "no está en cor
10						
11	España	Colau se propondrá para investirs	OKDIARIO	Barcelona en Comú prc	https://okdiario.com/espana,	['Barcelona en Comú propone que Ada Colau se presente co

Herramientas utilizadas

Whoosh

- Indexación.
- Búsqueda de noticias por diferentes campos.



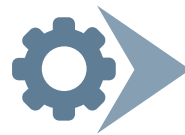
Herramientas utilizadas

Sistemas de recomendación

- Recomendar noticias
- Basado en contenido: Pesos [70% categoría, 30% autor]



Preferencias del usuario



Recomendaciones personalizadas



Herramientas utilizadas

Django

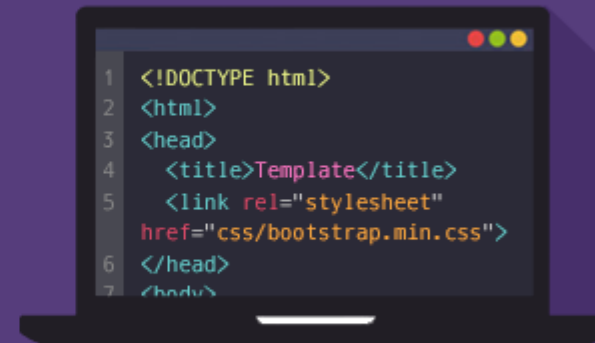
- Interfaz web.
- Punto de encuentro de todas las herramientas.

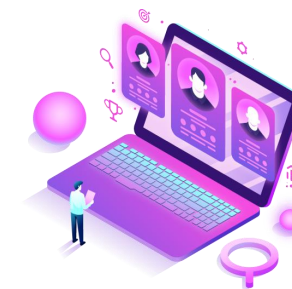


Herramientas utilizadas

Bootstrap

- Facilitar diseño.





Sistema implementado



Sistema implementado

Scraping realizado sobre OkDiario



A screenshot of the OkDiario website. The header features the 'okdiario' logo with the tagline 'EL SITIO DE LOS INCONFORMISTAS'. Below the logo is a navigation menu with links: ESPAÑA, ECONOMÍA, DEPORTES, INVESTIGACIÓN, INTERNACIONAL, OPINIÓN, LOOK, TELEVISIÓN, SOCIEDAD, and VÍDEOS. The main content area displays three news articles. The top article is titled 'HAN PASADO 271 DÍAS... Y LA DEMANDA DE SÁNCHEZ SIGUE SIN LLEGAR' and includes a calendar graphic for June 2019. The middle article is titled 'LLIBERTAT L'AUTODETERMINACIÓ ÉS UN DRET' and features a photo of a man. The bottom article is titled 'TORRA ORDENA A LA GOLPISTA ÒMNium QUE ORGANICE LA DESOBEDIENCIA CIVIL TRAS LA SENTENCIA DEL 1-0' and includes a photo of a man and a poster for 'JORNADAS SOBRE LA DESOBEDIENCIA CIVIL'.

Sistema implementado



Primero, desde el punto de vista del



Usuario

Login



okdiario

okdiario

Ingresa con tu cuenta

La actualidad te espera

¡Loguéate!

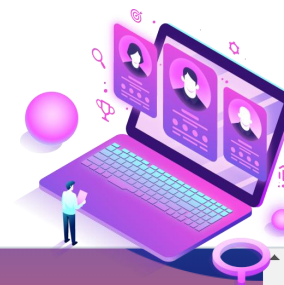
Nombre de usuario:

Contraseña:

LOGIN

REGISTRARSE

Registro



Consigue una cuenta
Es rápido y fácil

¡Regístrate!

Nombre de usuario:

Requerido. 150 caracteres como máximo. Únicamente letras, dígitos y @/./+/-/_

Contraseña:

- Su contraseña no puede asemejarse tanto a su otra información personal.
- Su contraseña debe contener al menos 8 caracteres.
- Su contraseña no puede ser una clave utilizada comunmente.
- Su contraseña no puede ser completamente numérica.

Contraseña (confirmación):

Para verificar, introduzca la misma contraseña anterior.

COMPLETAR REGISTRO

YA TENGO UNA CUENTA

Inicio



 Buscador basado en Whoosh Recomendador Cerrar sesión



Buscador y recomendador de noticias

Encuentra noticias que te interesen rápidamente

[BUSCAR NOTICIAS](#) [RECOMENDAR NOTICIAS](#)

Esta página está realizada con las siguientes herramientas



BeautifulSoup se utiliza para recopilar

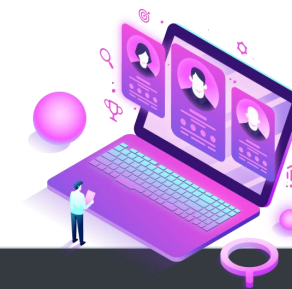


Whoosh ha sido utilizado para realizar la indexación y búsqueda de noticias de



Django se utiliza para la construcción de

Búsqueda



okdiario

Buscador basado en Whoosh

Recomendador

Cerrar sesión

okdiario

Buscador basado en Whoosh

Encuentre noticias rápidamente

BUSCAR NOTICIAS

RECOMENDAR NOTICIAS

Buscar Noticia por Términos

Términos:

CONSULTAR

Noticias encontradas



 Buscador basado en Whoosh Recomendador Cerrar sesión



Buscador basado en Whoosh

Encuentre noticias rápidamente

[BUSCAR NOTICIAS](#) [RECOMENDAR NOTICIAS](#)

Noticias encontradas



Seccion: Tribunas



Seccion: Andalucía

Noticias encontradas



Noticias encontradas



Seccion: Tribunas

Autor: Carlos Dávila

[Ir a la noticia original](#)

Descripcion: Que quede claro, muy claro. No tengo ninguna razón ni profesional, ni personal, ni política para hablar bien de Vox. Todavía creo conservar algún amigo en esa formación política,...

LEER MÁS



Seccion: España

Autor: Fernán González

[Ir a la noticia original](#)



Seccion: Andalucía

Autor: OKDIARIO

[Ir a la noticia original](#)

Descripcion: Macarena Olona, candidata de Vox al Congreso por Granada, visita la Plaza de Toros de Granada para estar con los alumnos de la Escuela Taurina de la ciudad. Así...

LEER MÁS

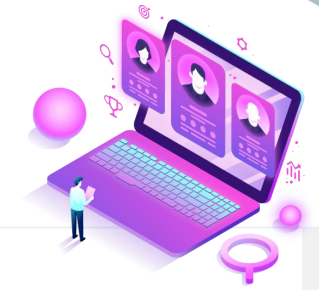


Seccion: Madrid

Autor: Fernán González

[Ir a la noticia original](#)

Detalles de noticia



VOLVER

Vox: razones y sinrazones (con El Yunque de fondo)

Categoría: Tribunas

Autor: Carlos Dávila

[Ir a la noticia original](#)

Que quede claro, muy claro. No tengo ninguna razón ni profesional, ni personal, ni política para hablar bien de Vox. Todavía creo conservar algún amigo en esa formación política, y gente con la que evidentemente no me tomaría ni un café. Y aquí no incluyo a Santi Abascal. Hablo, por ejemplo, de la gente de HazteOir, de esa peligrosísima secta pseudoreligiosa que se llama El Yunque. Son gente con la que no solo no me tomaría un café, sino que además la quiero combatir personalmente, ideológicamente y políticamente porque son nocivos para la sociedad, también para los principios que dicen defender. Dicho esto, afirmo, con igual rotundidad, que la gente de Vox tienen toda la razón para querer ser útiles y estar entre las personas y los partidos que van a negociar la sustitución en muchos sitios de España de toda la izquierda radical que nos ha venido malgobernando durante todo este tiempo. ¿Por qué Vox que respeta –o yo por lo menos creo que sí– el sistema constitucional no puede aparecer como socio de Ciudadanos o del propio Partido Popular y un partido de ultraizquierda comunista, madurista, marxista-leninista como Podemos puede estar como el perejil en todas las salsas? No termino de entenderlo ni lo voy a entender nunca. Vox tiene todo el derecho a estar en esas negociaciones, e incluso a pedir su cuota de responsabilidad. Otra cosa es que las negociaciones vayan por un lado o por otro. Pero aquí lo importante es lo que he dicho al principio. Tienen toda la razón para estar en las negociaciones. Y desde luego para ser críticos, pero también útiles, en la sustitución de la izquierda radical que nos ha venido utilizando como moneda de cambio para llegar a un Gobierno marxista-leninista como el que ha habido en muchos sitios de España hasta este momento.

Noticias recomendadas



Noticias recomendadas



Seccion: Tribunas

Autor: Carlos Dávila

[Ir a la noticia original](#)

Descripcion: Y no me refiero a este PSOE de ahora mismo y del inaudito e increíble Sánchez; no, no voy a denunciar -al menos hoy- las maniobras, siempre en la...

LEER MÁS



Seccion: Tribunas

Autor: Carlos Dávila



Seccion: Tribunas

Autor: Carlos Dávila

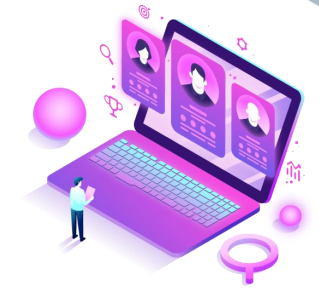
[Ir a la noticia original](#)

Descripcion: Que quede claro, muy claro. No tengo ninguna razón ni profesional, ni personal, ni política para hablar bien de Vox. Todavía creo conservar algún amigo en esa formación política,...

LEER MÁS

Mostrar iconos ocultos

Sistema implementado



Ahora, desde el punto de vista del



Administrador

Inicio



[Poblar base de datos](#) [Poblar índice de Whoosh](#) [Vista de administrador](#) [Cerrar sesión](#)



Buscador y recomendador de noticias

Encuentra noticias que te interesen rápidamente

Esta página está realizada con las siguientes herramientas



BeautifulSoup se utiliza para recopilar información en archivos HTML, haciendo uso de sus etiquetas.




Whoosh ha sido utilizado para realizar la indexación y búsqueda de noticias de manera eficiente.




Django se utiliza para la construcción de la web, de manera que se puedan realizar las búsquedas mediante una interfaz gráfica.

Poblar base de datos



 [Poblar base de datos](#) [Poblar índice de Whoosh](#) [Vista de administrador](#) [Cerrar sesión](#)


Base de datos cargada correctamente

Esta página está realizada con las siguientes herramientas



BeautifulSoup se utiliza para recopilar información en archivos HTML, haciendo uso de sus etiquetas.

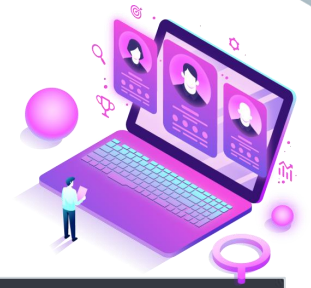


Whoosh ha sido utilizado para realizar la indexación y búsqueda de noticias de manera eficiente.



Django se utiliza para la construcción de la web, de manera que se puedan realizar las búsquedas mediante una interfaz gráfica.

Poblar índice de Whoosh



[Poblar base de datos](#) [Poblar índice de Whoosh](#) [Vista de administrador](#) [Cerrar sesión](#)



Índices de Whoosh cargados correctamente

Esta página está realizada con las siguientes herramientas



BeautifulSoup se utiliza para recopilar información en archivos HTML, haciendo uso de sus etiquetas.



Whoosh ha sido utilizado para realizar la indexación y búsqueda de noticias de manera eficiente.



Django se utiliza para la construcción de la web, de manera que se puedan realizar las búsquedas mediante una interfaz gráfica.

¿Preguntas?

