
Máster en Big Data y Business Analytics

2015 / 2016



SISTEMA DE GENERACIÓN DE ALERTAS DE BROTOS ALÉRGICOS POR GRAMÍNEAS

TRABAJO FIN DE MÁSTER

<https://othesoluciones.github.io/TFM/>

Autores

Sara Otero de Navascués Tomé
César Hernández Martín

Director:

Antonio Sarasa Cabezuelo

*Lo que parece ser el final,
suele ser en realidad,
un nuevo comienzo...*

CONTENIDO

INTRODUCCIÓN	1
OBJETIVO	3
CONFIGURACIÓN DEL PROYECTO	4
MODELO	11
BÚSQUEDA DE NOTICIAS	24
APLICACIÓN WEB	27
TRABAJOS FUTUROS	41
CONCLUSIÓN	42
BIBLIOGRAFÍA	43

INTRODUCCIÓN

Se estima que entre el 30 y el 40% de la población a nivel mundial se encuentra afectada por alguna enfermedad alérgica, el aumento de la temperatura y la contaminación atmosférica están provocando un incremento en la concentración del polen que se encuentra en la atmósfera y se espera que si no mejora la calidad del aire el porcentaje de la población afectada continúe aumentando.

Busquemos el origen del problema, el polen. El polen es el polvo, más o menos grueso, producido por los estambres de las flores, que se encarga de transportar las células espermáticas hasta el aparato reproductor femenino de otras flores. Hay que tener en cuenta, que una única planta es capaz de producir miles y miles de granos de polen. Aunque no todas polinizan al mismo tiempo, la mayoría lo hacen entre primavera y verano, por lo que el nivel de polen varía dependiendo de la planta y la época del año.

La sintomatología en los humanos provocada por la sensibilidad clínica a los alérgenos que se encuentran en el polen, se denomina polinosis. Siendo uno de los más comunes el de gramíneas. La familia de las gramíneas es una de las más numerosas, consta de casi 700 géneros y de unas 12.000 especies. Se calcula que las gramíneas cubren un 20% de la superficie del mundo.

No fue hasta 1873 cuando Charles Harrison Blackley asoció la causa de esta enfermedad al nivel de polen existente en el aire. A partir de entonces, se ha intentado predecir los lugares y épocas del año de mayor virulencia, al principio teniendo en cuenta únicamente las zonas cercanas a plantaciones y los meses de polinización, pero actualmente, y gracias a las técnicas Big Data, disponemos de la capacidad de analizar otros factores que también influyen en el nivel de alergia como puede ser la calidad del aire, la temperatura, la probabilidad de lluvia, etc.

En este proyecto hemos desarrollado una aplicación que ejecuta un modelo que trata de realizar dichas predicciones para los municipios de la Comunidad de Madrid. La aplicación esta accesible al público en general a través de la página Web:

<http://gramineas-madrid.herokuapp.com>

Los datos que se muestran en la aplicación, y sobre los que se realiza el modelo, son actualizados de forma diaria mediante un proceso batch que se ejecuta a primera hora de la mañana. Aun así, el nivel de alerta se actualiza, tanto en base de datos como en la aplicación, cada vez que un usuario envía un reporte. Los reportes se realizan a nivel de municipio, al asignarle al usuario el nivel de alerta que quiera notificar.

Además de la recogida de datos, en este batch se ejecuta el algoritmo que realiza la predicción, el motor de búsqueda que actualiza las noticias que se muestran por pantalla y el envío automático de correos electrónicos a aquellos usuarios que deseen recibir notificaciones sobre un determinado municipio.

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

La intención de esta aplicación es ofrecer a los usuarios un servicio que les proporcione una orientación sobre el nivel de polen de los Municipios de la Comunidad de Madrid. Queremos resaltar que los datos, sobre los que se realiza la predicción inicial, proporcionada a primera hora de la mañana, se nutren de fuentes oficiales. Mientras que la que se muestra a lo largo del día puede depender de los reportes de alertas realizados por los usuarios de la aplicación.

A lo largo de este trabajo detallaremos como se ha realizado, tanto el algoritmo predictivo como la Web. Y se indicará los usos que se puede hacer de la aplicación.

Por último, con el objeto de centralizar todo el contenido de este trabajo de Fin de Máster, hemos creado una página web desde la cual podemos descargar la aplicación desde su repositorio de GitHub, así como esta memoria y la presentación que se mostrará en el acto de lectura del TFM del 30 de septiembre de 2016:

<https://othesoluciones.github.io/TFM/>



Máster en Big Data y Business Analytics 2015/2016

OBJETIVO

El objetivo de este proyecto consiste en el diseño, desarrollo e implementación de una Aplicación Web dirigida al estudio y predicción de brotes alérgicos producidos por el polen de gramíneas.

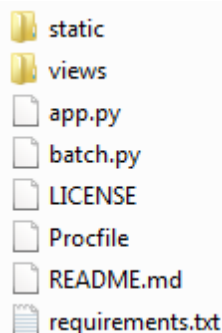
El sistema consistirá en una prueba de concepto para una única zona geográfica, la Comunidad de Madrid, implementándose de forma que se pueda ser extensible a cualquier zona y tipo de polen.

Para la realización de dicho proyecto, se utilizarán los conceptos y conocimientos adquiridos a lo largo del curso así como nuevas competencias que han sido necesarias para la realización del presente trabajo.

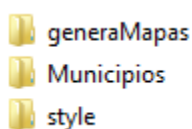
CONFIGURACIÓN DEL PROYECTO

Hemos decidido abordar la construcción del sistema de generación de alertas de brotes alérgicos por gramíneas como una aplicación web, quedando finalmente con la siguiente configuración:

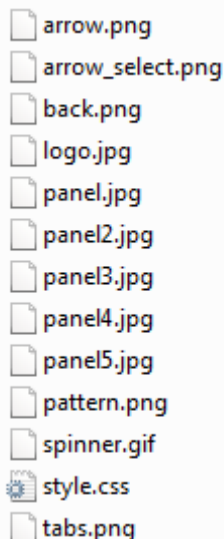
- Directorio Raíz del Proyecto:



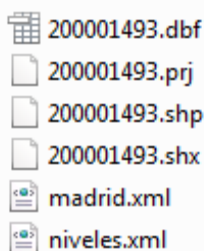
- Directorio **static**



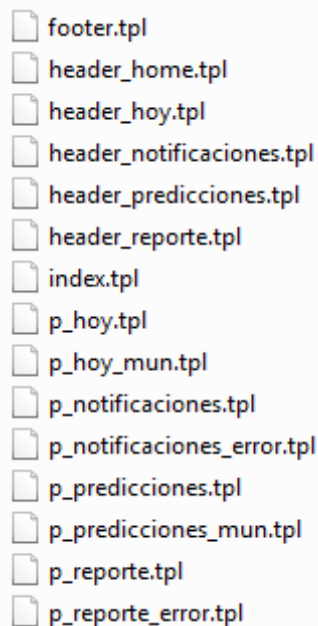
- Directorio **static/style**



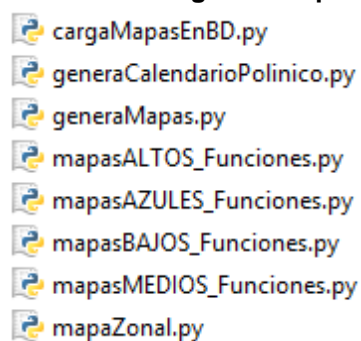
- Directorio **static/Municipios**



- Directorio **views**



- Directorio **static/generaMapas**



Para lograr esta estructura vamos a comentar las tecnologías que hemos utilizado en el desarrollo de la aplicación.

Control de Versiones



Como sistema de control de versiones, hemos optado por el distribuido: **Git** utilizando para la gestión del proyecto la plataforma de desarrollo colaborativo de software, **GitHub**.

Hemos alojado en un repositorio de esta plataforma el código que hemos desarrollado durante la elaboración de este proyecto, que podrá ser distribuido bajo licencia del MIT, por lo cual admitimos el uso comercial, la redistribución y la modificación del software con la condición de proporcionar una copia de la licencia con el software que hemos distribuido.

Además hemos elegido GitHub porque permite almacenar proyectos en la mayoría de lenguajes de programación, ofreciendo también la posibilidad de conectarse con plataformas PaaS (Platform as Service) como OpenShift o Heroku para automatizar el proceso del despliegue.

Por último, esta es la URL de nuestro repositorio GitHub:

<https://github.com/othsoluciones/TFM>

Lenguaje de programación



Python es un lenguaje de programación multiparadigma, lo que significa que admite varios estilos de programación: orientada a objetos, funcional e imperativa. Se pueden utilizar otro tipo de paradigmas mediante el uso de extensiones.

Fue creado por [Guido van Rossum](#), Países Bajos, a finales de la década de los ochenta como un sucesor del lenguaje ABC. Y recibe su nombre en honor a los humoristas británicos Monty Python.

Se caracteriza por ser un lenguaje interpretado de tipado dinámico y multiplataforma. Su sintaxis está diseñada para una fácil lectura (los operadores lógicos como `and` u `or`, se codifican mediante palabras en lugar de símbolos) y los bloques de código se delimitan mediante tabulaciones o espacios, en lugar de utilizar llaves o paréntesis.

Además de estas características, lo que nos ha hecho decantarnos por este lenguaje de programación es que ha sido la herramienta que más hemos utilizado durante todo el curso para el análisis de datos, y hemos querido dar continuidad a los conocimientos aprendidos durante el año en la realización del proyecto.

Otros aspectos importantes en la elección de Python es que se trata de un proyecto desarrollado en código abierto y cuenta con una gran comunidad de usuarios activa y en continuo crecimiento, esto hace que dispongamos de una importante cantidad de biblioteca de librerías que pueden ser utilizadas para múltiples tareas. Así en este proyecto hemos utilizado las siguientes:

- **Numpy:** es uno de los principales paquetes científicos de Python. Permite la manipulación de matrices multidimensionales a gran velocidad.
- **Pandas:** librería que está considerada como una de las herramientas existentes más flexibles y potentes para la manipulación y análisis de información gracias a la estructura de datos que proporciona.
- **Matplotlib:** es una librería de Python que permite realizar de forma sencilla y en pocas líneas de código multitud de gráficos.
- **Pyshp:** paquete que nos permite el procesamiento de ficheros shp para la generación de los mapas que se van a mostrar en la aplicación mediante Matplotlib.
- **Xmltodict:** módulo de Python que facilita la lectura de XML, traduciéndolo a lenguaje JSON.
- **Lxml:** librería que se utiliza para el procesamiento de XML.
- **Beautifulsoup:** es una de las bibliotecas más utilizadas para la extracción de

información de páginas web (web scrapping) facilitando la búsqueda, iteración y modificación del árbol generado por el código HTML.

- **Google-api-python-client:** librería que permite realizar búsquedas en la Web mediante la creación un motor de búsqueda.
- **Pymongo:** es la herramienta utilizada para interactuar con bases de datos **MongoDB**, que tal y como contaremos a continuación es la base de datos que hemos elegido en esta aplicación.
- **APScheduler:** librería de Python, cuyo nombre completo es Advanced Python Scheduler, que posibilita la ejecución del código a futuro, programando la hora de ejecución tanto de forma periódica como una única vez. Gracias su inclusión ejecutamos de forma escalonada todas las funciones definidas en nuestro archivo **batch.py**

Además Python ofrece la posibilidad de implementar aplicaciones web mediante frameworks, como Django, Pyramid Flask o **Bottle**,

Todas estas librerías se encuentran definidas, junto con su versión, en el fichero **requirements.txt**

Framework



Entre los frameworks anteriormente citados, finalmente nos hemos decantado por desarrollar la aplicación con **Bottle**, ya que se trata de un rápido, simple y ligero WSGI micro-framework para Python, además no tiene más dependencias externas aparte de las librerías estándar de Python y se distribuye en un único fichero *.py (en nuestra aplicación se corresponde con el fichero **app.py**). Entre sus funciones podemos destacar que:

- Bottle permite redigirir a URLs específicas desde funciones
- Bottle dispone de un motor propio de plantillas y ofrece soporte para otros, como Jinja2, Mako o Cheetah (en nuestro proyecto todas las plantillas que utilizamos están almacenada en el directorio **views**). En estas plantillas, además, se permite el uso de librerías externas de Javascript como por ejemplo **JQuery** que nos ha permitido de manera sencilla manipular documentos HTML, crear animaciones y gestionar eventos
- Bottle ofrece acceso entre otros a datos de formularios, subida de ficheros, cookies o cabeceras.
- El servidor HTTP de Bottle no es únicamente para desarrollo interno, sino que además es un servidor de soporte para otros, como fapws3, bjoern, o incluso para PaaS como Google App Engine, OpenShift o **Heroku**

PaaS (Platform as a Service)



Es una plataforma que proporciona un servicio de computación en la Nube (PaaS, Platform as a Service con las siglas en inglés), que permite desarrollar y desplegar aplicaciones de manera sencilla. Admite multitud de lenguajes de programación como: Node, Ruby, Java, PHP, **Python**, Go, Scala, o Clojure.

Las principales razones por las que hemos elegido Heroku como PaaS son:

- Se puede habilitar la conexión con el repositorio de **GitHub** donde se almacena el software
- Se puede automatizar el proceso de despliegue simplemente realizando una actualización del software en **GitHub**, por lo cual podemos tener desplegada nuestra aplicación en Producción tras la realización de un cambio en menos de dos minutos (también se permite el despliegue de forma manual desde el Dashboard de Heroku e incluso desde línea de comandos).
- Fácil acceso a los logs de la aplicación desde el Dashboard o descarga en local desde línea de comandos.
- Heroku ofrece de manera gratuita 512 MB de RAM para la aplicación, y la posibilidad de definir sobre ella hasta dos procesos distintos llamados Dynos: uno tipo web (nuestra aplicación Python-Bottle **app.py**) y uno tipo clock (el fichero Python **batch.py**). Configurándose esto de forma muy sencilla en el fichero **Procfile**:

web: python app.py

clock: python batch.py

- Heroku ofrece más de 100 add-ons que pueden ser añadidas en las aplicaciones que aloja, en nuestro caso hemos utilizado uno: **Heroku Scheduler** que nos permite ejecutar de forma diaria y programada nuestro fichero de actualización del sistema **batch.py**

Base de Datos



Para el almacenamiento de los datos necesarios para el correcto funcionamiento de nuestra aplicación hemos optado por una Base de Datos No SQL de tipo documental: **MongoDB**, que está orientada a documentos y almacena la información en formato BSON. Además es multiplataforma y está disponible para los sistemas operativos Windows, Linux, OS X y Solaris.

MongoDB prima la disponibilidad y la consistencia frente a la disponibilidad en presencia de particiones (Teorema CAP).



En 2007 comenzó su desarrollo por la empresa de software 10gen Inc. En 2009 se publicó de forma independiente con licencia de código abierto, aunque no fue hasta 2011 cuando se consideró lista para usarse en producción.

Hemos elegido esta base de datos debido a que nuestra fuente de datos está fundamentada principalmente en web-scraping, y por lo tanto no podemos estar seguros de la estructura y existencia de los datos recogidos con esta técnica, lo que se encaja en una base de datos sin esquema.

Por otro lado, hemos escogido MongoDB porque ha sido una de las bases de datos con las que hemos trabajado a lo largo del curso, y por ello hemos querido afianzar los conocimientos adquiridos.

Otra razón importante ha sido la existencia de la librería **PyMongo** para Python, que permite de forma sencilla la integración de esta base de datos.

Además, MongoDB admite el almacenamiento de datos de diferentes naturaleza y tamaño, como pueden ser datos de formato textual o en formato imagen, como los mapas de los municipios que se muestran en la aplicación.

Teniendo en cuenta que esta aplicación se trata de una prueba de concepto y se espera que en un futuro se puedan incluir nuevas variables, tipos de pólenes o localizaciones, se desea que la base de datos se fácilmente escalable de horizontalmente, una característica importante de MongoDB.



Por último, hemos decidido utilizar **MLab** anteriormente conocido como MongoLab, como servicio SaaS (Software as a Service) de MongoDB en la Nube para poder tener acceso a los datos en el entorno productivo.

Esta compañía fue creada en 2011, y gestiona actualmente más de 350.000 bases de datos a través de tres principales proveedores de la nube en 25 centros de datos en todo el mundo.

Entre los distintos planes que ofrece MLab, nos hemos decantado por el gratuito, ya que pone a nuestra disposición 500 MB para el almacenamiento de datos, que ha sido suficiente para la propuesta actual de nuestra aplicación.

MODELO

Antes de comenzar a describir el algoritmo utilizado para realizar la predicción de brotes alérgicos, es necesario conocer cómo y cuándo se producen las reacciones alérgicas generadas por el polen de gramíneas.

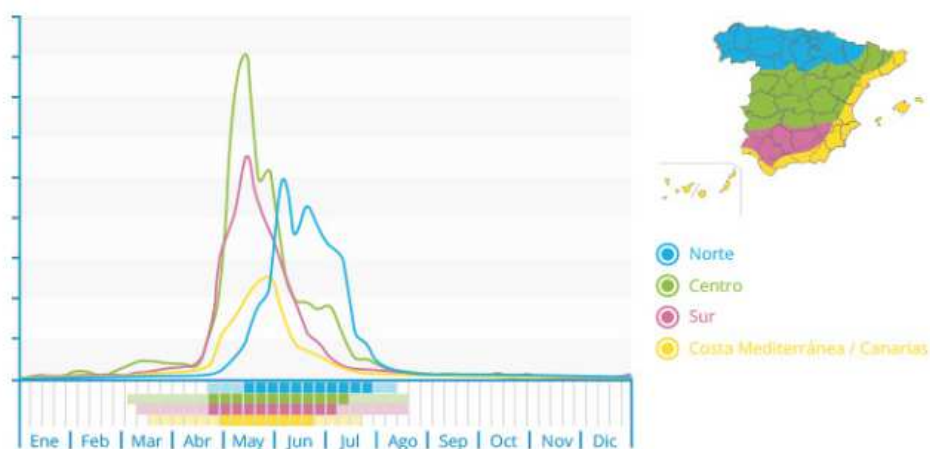
Las gramíneas o poáceas son una familia de plantas del orden de las monocotiledóneas, a la que pertenecen un gran número de especies usados tanto para el consumo humano (caña de azúcar, trigo, arroz, maíz...) como para el animal (pastos, granos...).



<http://www.plantarteentuoasis.com/2014/04/gramineas-jardin-texturas-composiciones.html>

En 1997, el Anuario del Ministerio de Agricultura, Pesca y Alimentación estimaba en 95.959 hectáreas la superficie de la Comunidad de Madrid destinada al cultivo de cereales. Otras especies crecen de forma espontánea en parques y jardines o al borde de los cultivos.

En la Comunidad de Madrid se considera que el polen generado por las gramíneas es la principal causa de la polinosis, al igual que en el resto de España. Este polen se encuentra en el aire prácticamente durante todo el año, sin embargo los meses de mayor incidencia quedan comprendidos de marzo a julio, destacando principalmente las semanas 21, 22 y 23.



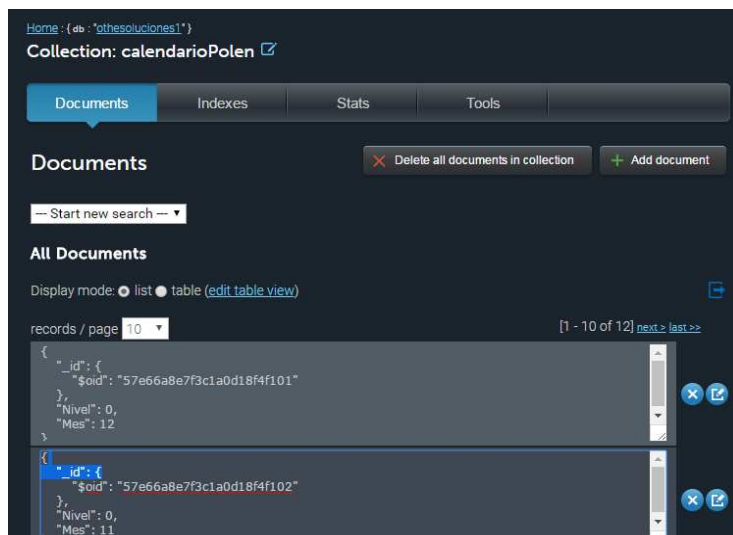
Hemos partido del calendario polínico como base para realizar la predicción, ya que el nivel de polen es claramente un dato estacional, para lo cual hemos establecido tres niveles distintos de alerta (Bajo, Medio y Alto). Clasificando de forma inicial cada mes en uno de estos niveles:

Niveles de partida

MES	NIVEL
Enero	Bajo
Febrero	Bajo
Marzo	Medio
Abril	Alto
Mayo	Alto
Junio	Alto
Julio	Medio
Agosto	Bajo
Septiembre	Bajo
Octubre	Bajo
Noviembre	Bajo
Diciembre	Bajo

Cada uno de estos niveles tiene asociado un valor numérico, que se irá incrementando o reduciendo dependiendo del resto de factores incluidos en el algoritmo. De esta forma al nivel Bajo se le asocia el valor 0, al Medio el 1 y al Alto el 2.

Para tratar esta información de una forma más accesible y cómoda hemos creado en nuestra base de datos MongoDB una colección llamada **calendarioPolen**, generada al ejecutar el archivo **generaCalendarioPolinico.py**

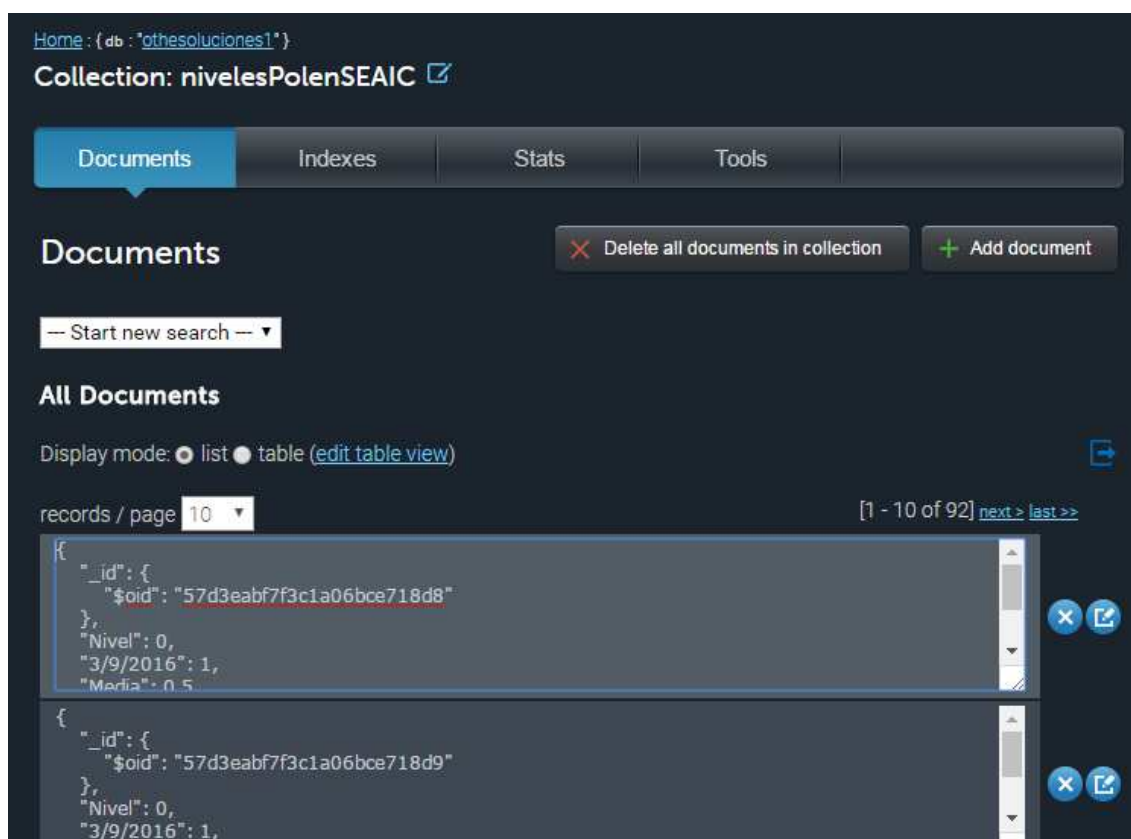


Posteriormente, se ha tenido en cuenta el nivel medio diario de granos de polen por metro cúbico de aire, que nos proporciona la *SEAIC (Sociedad Española de Alergología e Inmunología Clínica)* para la semana anterior, con este dato realizamos la media de la semana y le asociamos un valor, si es menos a 200 gr/m³ se le asigna el valor 0, si se encuentra entre 200 y 1000 gr/m³ el valor 1 y para cantidades superiores asignamos el valor 2. De esta forma si el nivel proporcionado por el calendario polínico es superior al asignado teniendo en cuenta la concentración de polen en el aire, incrementamos el primero en un 0,5, mientras que si es menor disminuimos el nivel calculado en la misma proporción. De esta forma vamos a ir obteniendo un valor numérico que finalmente traduciremos en el nivel de alerta calculado por nuestro sistema.

Los datos recuperados de la página de la SEAIC se van almacenando en la colección **nivelesPolenSEAIC**, especificando la media de la semana, la semana en la que os encontramos, los valores por día, el tipo de polen y el nivel calculado para dicha media:

```
{
  "_id": {
    "$oid": "57d3eabf7f3c1a06bce718d8"
  },
  "Nivel": 0,
  "3/9/2016": 1,
  "Media": 0.5,
  "Semana": 36,
  "Polen/Fecha": "Gramineas",
  "4/9/2016": 0
}
```


Actualmente ya estamos recuperando datos para otros tipos de polen distinto de Gramíneas para poder ampliar fácilmente la aplicación en un futuro.



Vamos a continuar con otros factores que también influyen en la virulencia de los brotes, como la calidad del aire.

Varios estudios han demostrado que la contaminación, sobre todo la producida por las partículas emitidas por los motores de los coches diésel, modifica la naturaleza de los pólenes volviéndolos más agresivos. Según la SEAIC, los distintos contaminantes terminan depositándose en el suelo, afectando al desarrollo natural de semillas, raíces y plantas, alterando sus características fisiológicas y convirtiendo a los pólenes en más alérgicos y potentes. Además la contaminación afecta a las vías respiratorias irritando las mucosas de nariz, faringe y pulmones, lo que agrava la sintomatología propia de la alergia. Esto explica que haya más alergias en las ciudades que en el campo, aunque los niveles de polen sean inferiores. De hecho, el aumento en el número de pacientes sensibilizados a este tipo de polen ha sufrido un ascenso que ha coincidido con el progresivo aumento de número de coches diésel, actualmente alrededor del 60% de los vehículos que circulan por Madrid son diésel, mientras que hace veinte años era del 25%.



Antonio Valero, presidente de la SCAIC (Sociedad Catalana de Alergia e Inmunología Clínica) explica que el polen, en áreas de alta contaminación como son las grandes ciudades (Madrid, Barcelona,...), contiene mayor cantidad de proteínas definidas como alérgicas. En concreto, las partículas emitidas por motores diésel en contacto con el polen pueden ocasionar el desgrane de la planta, facilitando la llegada de estas partículas a las vías respiratorias (otro aspecto a tener en cuenta es que los motores diésel pueden emitir hasta 100 veces más partículas que los de gasolina). El cambio climático también afecta al calendario polínico de las plantas ya que avanza y alarga el periodo de polinización.

Habiendo contrastado que las emisiones producidas por motores diésel, así como la polución en general, agravan la sintomatología de los procesos alérgicos, veamos cuales son los contaminantes causantes de este empeoramiento y cuya concentración se ha tenido en cuenta a la hora de realizar las predicciones:

Dióxido de nitrógeno (NO₂): se forma como subproducto en los procesos de combustión a altas temperaturas, como en los vehículos motorizados y las plantas eléctricas. Por ello es un contaminante frecuente en zonas urbanas. Y el causante de que en Madrid se obligase a reducir la velocidad máxima de los vehículos en determinadas vías. Causa efectos nocivos en la salud, especialmente en el sistema respiratorio.

Partículas en suspensión < PM₁₀: partículas sólidas de 10 milésimas de milímetros de tamaño. Estas partículas debido a su tamaño no pasan los filtros corporales quedándose en la nariz y garganta. La combustión de carburantes fósiles es una de las principales fuentes de estas partículas en suspensión en las ciudades.

Partículas en suspensión < PM_{2,5}: partículas sólidas de 2,5 milésimas de milímetros de tamaño. Son partículas tan pequeñas que atraviesan los tejidos corporales depositándose en los bronquiolos, de donde no pueden extraerse. En gran proporción provienen de las emisiones de los vehículos diésel.

Monóxido de carbono (CO): producto intermedio de la oxidación de un hidrocarburo, es un gas incoloro altamente tóxico, que puede producir la muerte al inhalarlo en niveles

elevados. Se produce por la combustión deficiente de sustancias como gasolina, keroseno, carbón, petróleo, tabaco o madera. Los vehículos con el motor encendido, especialmente los de gasolina, también lo despiden.

Concentración de Ozono (O3): en las capas más altas de la atmósfera, el ozono sirve como filtro solar y nos protege de los niveles elevados de radiación ultravioleta procedente del sol. Sin embargo, al nivel del suelo puede ser perjudicial. Su exposición puede aumentar la susceptibilidad a los alérgenos respiratorios. Los niveles de ozono suelen ser menores en zonas urbanas contaminadas ya que reacciona con otros contaminantes como el NO.

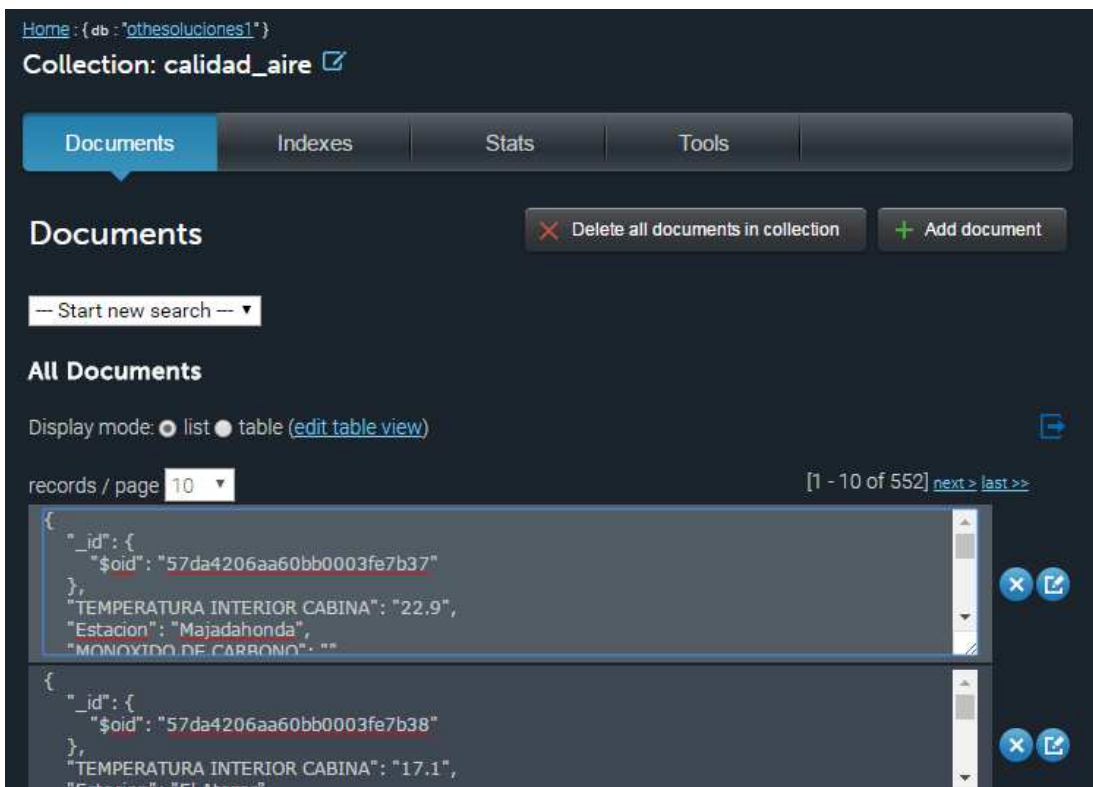
Dióxido de azufre (SO2): gas incoloro, irritante, de olor penetrante. Se considera uno de los responsables de la lluvia ácida. Su exposición comporta graves riesgos para la salud, ya que pasa directamente al sistema circulatorio a través de las vías respiratorias. La principal fuente de emisión es la combustión de productos petrolíferos y la quema de carbón en centrales eléctricas y calefacciones.

Monóxido de nitrógeno (NO): gas incoloro y difícilmente soluble en agua. Constituye uno de los contaminantes de la atmósfera que forma parte de la lluvia ácida. Es generado, entre otros, por los motores de explosión de los coches y expulsado a la atmósfera mediante los tubos de escape.

Estos datos se recuperan de la página de calidad del aire de la Comunidad de Madrid que se indica en la bibliografía y son almacenados en la colección **calidad_aire** con la siguiente estructura:

```
{
  "_id": {
    "$oid": "57da4206aa60bb0003fe7b37"
  },
  "TEMPERATURA INTERIOR CABINA": "22.9",
  "Estacion": "Majadahonda",
  "MONOXIDO DE CARBONO": "",
  "PRESION BAROMETRICA": "931",
  "CONCENTRACION DE TOLUENO": "",
  "HUMEDAD RELATIVA": "92",
  "VELOCIDAD DEL VIENTO": "0.4",
  "Hora": "06:00",
  "CONCENTRACION DE OZONO": "9",
  "TEMPERATURA": "7.9",
  "CONCENTRACION DE XILENO": "",
  "PARTICULAS EN SUSPENSION < PM10": "14",
  "PARTICULAS EN SUSPENSION < PM2,5": "",
  "PRECIPITACION": "0.0",
  "DIOXIDO DE AZUFRE": "",
  "Zona": 4,
  "RADIACION SOLAR": "2",
  "DIRECCION DEL VIENTO": "91",
  "Dia": "15/09/2016",
  "DIOXIDO DE NITROGENO": "27",
  "CONCENTRACION DE BENCENO": "",
  "MONOXIDO DE NITROGENO": "6",
  "CONCENTRACION DE HIDROCARBUROS NO METAN": "",
  "CONCENTRACION DE HIDROCARBUROS TOTALES": ""
}
```

Para cada una de las localidades que poseen un captador de calidad de aire.



Home : {db : "othersoluciones1"}

Collection: **calidad_aire**

Documents Indexes Stats Tools

Documents Delete all documents in collection Add document

— Start new search —

All Documents

Display mode: ☒ list ☐ table ([edit table view](#))

records / page 10 [1 - 10 of 552] [next >](#) [last >>](#)

```
{
  "_id": {
    "$oid": "57da4206aa60bb0003fe7b37"
  },
  "TEMPERATURA INTERIOR CABINA": "22.9",
  "Estacion": "Majadahonda",
  "MONOXIDO DE CARBONO": ""
}
```

```
{
  "_id": {
    "$oid": "57da4206aa60bb0003fe7b38"
  },
  "TEMPERATURA INTERIOR CABINA": "17.1",
  "Estacion": "El Atazar"
}
```

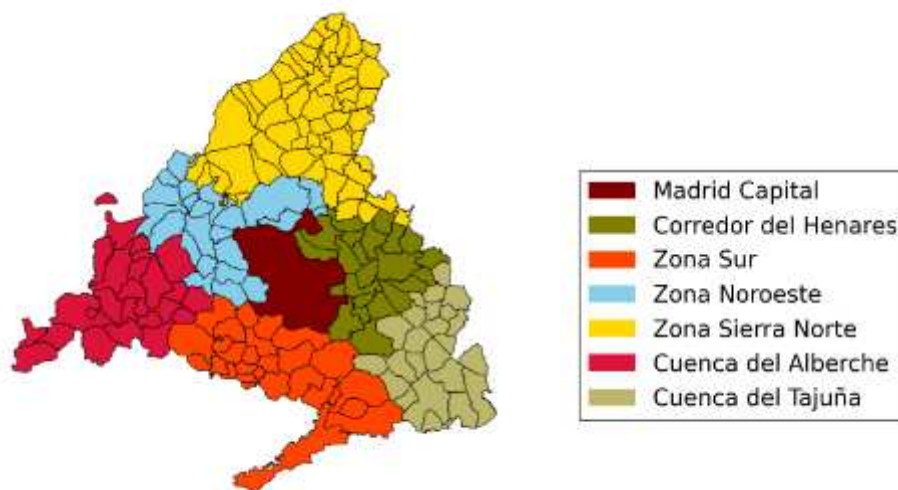
Se ha añadido 0,1 al nivel que teníamos previamente calculado cada vez que alguno de los contaminantes anteriormente descritos sobrepasaban unos valores límites. Estos valores límites son los que establece la legislación vigente sobre la calidad del aire y han sido recuperados de la siguiente tabla:

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Tabla resumen de legislación en materia de calidad del aire Valores límite y objetivos				
Real Decreto 102/2011				
Contaminante	Objeto de protección	Período de análisis	Valor	Categoría
Dióxido de nitrógeno (NO ₂)	Salud	Media anual	40 µg/m ³	Valor límite; En vigor desde 2010
	Salud	Media horaria; no podrán superarse en más de 18 ocasiones por año civil	200 µg/m ³	Valor límite; En vigor desde 2010
Óxidos de nitrógeno (NO _x)	Vegetación	Media anual	30 µg/m ³	Nivel crítico ⁽¹⁾ ; En vigor desde 2008
Partículas PM ₁₀	Salud	Media anual	40 µg/m ³	Valor límite; En vigor desde 2005
	Salud	Media diaria; no podrán superarse en más de 35 ocasiones por año	50 µg/m ³	Valor límite; En vigor desde 2005
Partículas PM _{2,5}	Salud	Media anual	25 µg/m ³	Valor objetivo, deberá ser alcanzado en 2010; valor límite entra en vigor en 2015
Ozono (O ₃)	Salud	Máxima diaria de las medias móviles octohorarias; no podrá superarse en más de 25 días por cada año civil de promedio en un periodo de 3 años	120 µg/m ³	Valor objetivo, deberá ser alcanzado en 2010 ⁽³⁾ (media años 2010, 2011, 2012)
	Vegetación	AOT40, calculado a partir de medias horarias de mayo a julio	18 000 µg/m ³ × h de promedio en un periodo de 5 años	Valor objetivo, deberá ser alcanzado en 2010 ⁽³⁾ (media años 2010, 2011, 2012, 2013 y 2014)
Dióxido de azufre (SO ₂)	Salud	Media horaria	350 µg/m ³	Valor límite; En vigor desde 2005
	Salud	Media diaria	125 µg/m ³	Valor límite; En vigor desde 2005
	Vegetación	Media anual e invierno (del 1 de octubre al 31 de marzo)	20 µg/m ³	Nivel crítico ⁽¹⁾ ; En vigor desde 2008
Monóxido de carbono (CO)	Salud	Máxima diaria de las medias móviles octohorarias	10 mg/m ³	Valor límite; En vigor desde 2005

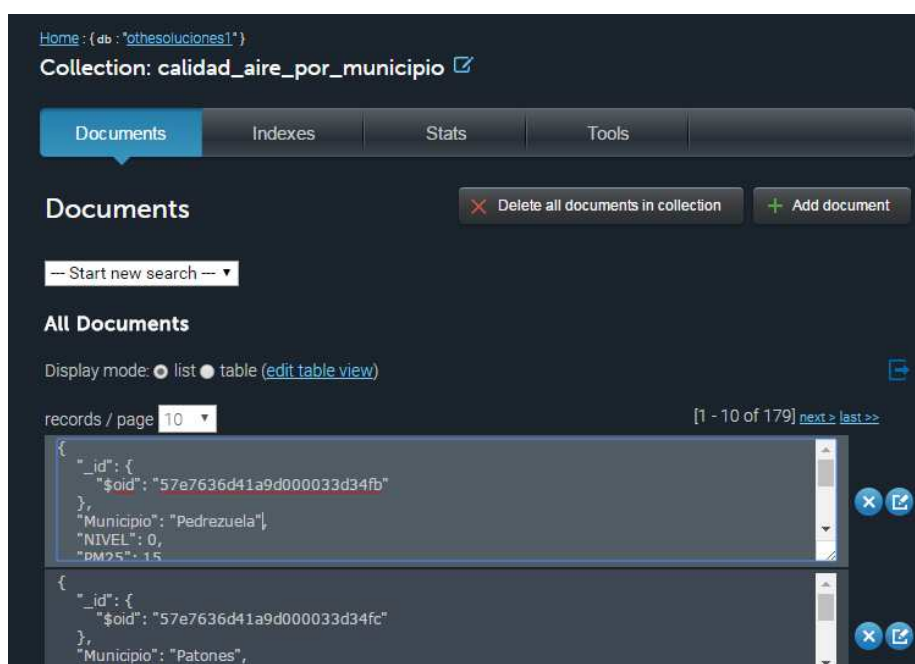
http://gestiona.madrid.org/azul_internet/html/web/2_3.htm?ESTADO_MENU=2_3

Para poder asociar a cada municipio el nivel calculado hemos tenido que dividir la Comunidad de Madrid en 7 zonas, ya que no existen captadores de calidad del aire en todas las localidades.



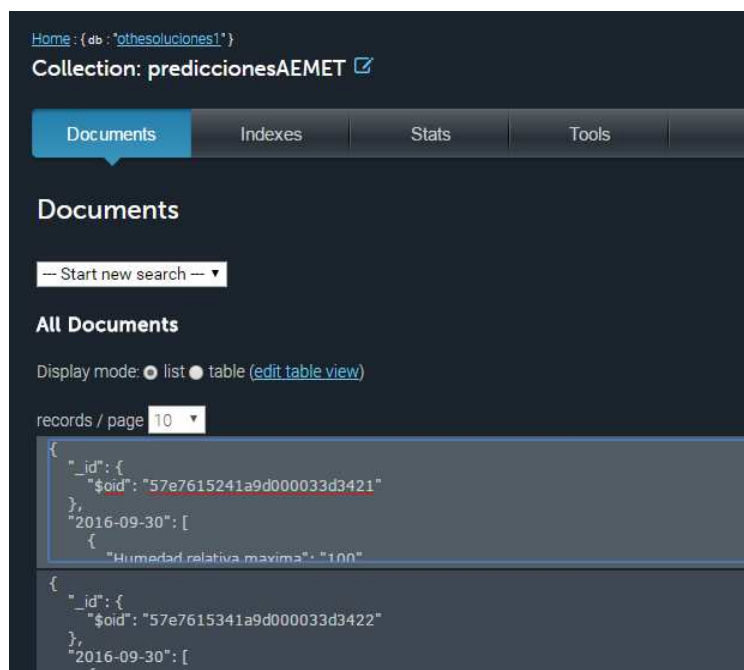
Se ha procedido recogiendo los valores registrados para cada contaminante en los captadores de una misma zona y extrapolando su media a todos los municipios de esa área. De esta forma ya tenemos diferenciado un nivel propio de cada población al que le añadiremos a continuación las predicciones meteorológicas realizadas por el AEMET (Agencia Estatal de Meteorología). Esta información se almacena en la colección **calidad_aire_por_municipio**:

```
{
  "_id": {
    "$oid": "57e7636d41a9d000033d34fb"
  },
  "Municipio": "Pedrezuela",
  "NIVEL": 0,
  "PM25": 15,
  "NO": 1,
  "O3": 84.5,
  "NO2": 4,
  "SO2": 2,
  "Zona": 5,
  "CO": 0.1,
  "PM10": 15.5
}
```



El clima también influyen en las alergias al polen, en los días lluviosos y los días nublados y sin viento los síntomas disminuyen, ya que la lluvia limpia la atmósfera y la ausencia de viento evita el desplazamiento del polen. Por el contrario, el tiempo cálido, seco y con viento favorece una mayor distribución del polen y, en consecuencia, un aumento de los síntomas.

Por ello hemos recuperado los datos meteorológicos de la página del AEMET para cada municipio y almacenados en la colección **prediccionesAEMET** de MongoDB:



Hemos tenido en cuenta los siguientes datos para los próximos tres días en cada uno de los distintos municipios de la Comunidad de Madrid:

- **Velocidad del viento (km/h):** Si la predicción de velocidad del viento da un valor superior a los 30km/h añadimos 0,3 al nivel de alerta calculado. Ya que el viento favorece el desplazamiento del polen.
- **Probabilidad de precipitaciones:** Si el AEMET predice un valor mayor al 30% le restamos a nuestro nivel 0,2 . La lluvia como hemos comentado disminuye los efectos, al limpiar el aire.
- **Humedad relativa mínima:** Con un valor superior al 30% disminuimos en 0,1 nuestra predicción.
- **Humedad relativa máxima:** Con un valor superior al 70% disminuimos en 0,1 el nivel calculado, ya que al haber mayor humedad se reducen los síntomas de la alergia tales como la congestión nasal y los picores.
- **Temperatura mínima:** Las concentraciones de polen suelen ser superiores en días cálidos. Por lo que con una temperatura mínima superior a 20°C añadimos 0,1 a nuestro nivel.
- **Temperatura máxima:** Y con una temperatura superior a los 30°C también le añadimos 0,1.

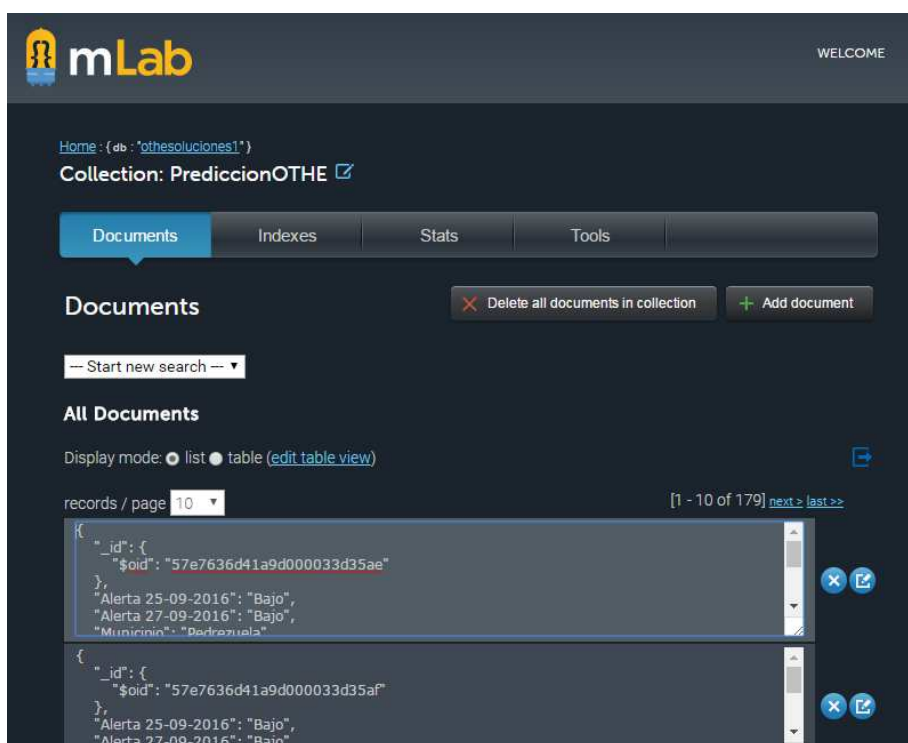
SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Con este cálculo obtenemos un valor numérico para cada municipio y para cada uno de los tres próximos días, este valor es redondeado y se le asigna un nivel de alerta, calculamos de esta forma la predicción inicial que se muestra en la aplicación:

Relación entre valor y nivel

VALOR CALCULADO	NIVEL OBTENIDO
<1	BAJO
>1 y <2	MEDIO
>2	ALTO

Estos datos son almacenados en la colección **PrediccionOTHE** de MongoDB



SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Donde cada municipio dispone de un documento con la siguiente información:

```
{
  "_id": {
    "$oid": "57e7636d41a9d000033d35ae"
  },
  "Alerta 25-09-2016": "Bajo",
  "Alerta 27-09-2016": "Bajo",
  "Municipio": "Pedrezuela",
  "Nivel 26-09-2016": 0,
  "Codigo": "28108",
  "Nivel 25-09-2016": 0,
  "Alerta 26-09-2016": "Bajo",
  "MunicipioOrig": "Pedrezuela",
  "Nivel 27-09-2016": 0
}
```

Ya tenemos por tanto la predicción que se muestra a primera hora de la mañana, pero a lo largo del día esta puede variar dependiendo de los reportes que haga el usuario a través de la web. Cada vez que un usuario realiza un reporte se vuelve a calcular el nivel de alerta para ese día, y ello se hace mediante la siguiente fórmula:

$$\text{nuevoNivel} = \text{nivelViejo} + (\text{nivel_alerta} - \text{round}(\text{nivelViejo})) * 0.45$$

Donde nivelViejo es el nivel existente hasta el momento en la colección de **PredicciónOTHE** para ese día y ese municipio y nivel_alerta es el proporcionado por el usuario en el reporte. De esta forma si se reporta un nivel inmediatamente superior al calculado inicialmente serán necesarios dos reportes para pasar a dicho nivel superior, pero si el nivel reportado es dos veces superior (nivel inicial: Bajo, nivel reportado: Alto) se actualizará al nivel intermedio. Actuando análogamente en el caso de que se reporten niveles inferiores. Una vez calculado el nuevo nivel este valor será actualizado en la colección para ser mostrado a partir de ese momento por pantalla, de forma que los reportes se actualizan en tiempo real.

Para hacer una prueba actualizamos el nivel de Pedrezuela a Alto

Repórtanos alertas

Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Municipio	Pedrezuela
Nivel	Alto
Reportar alerta	

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Repórtanos alertas

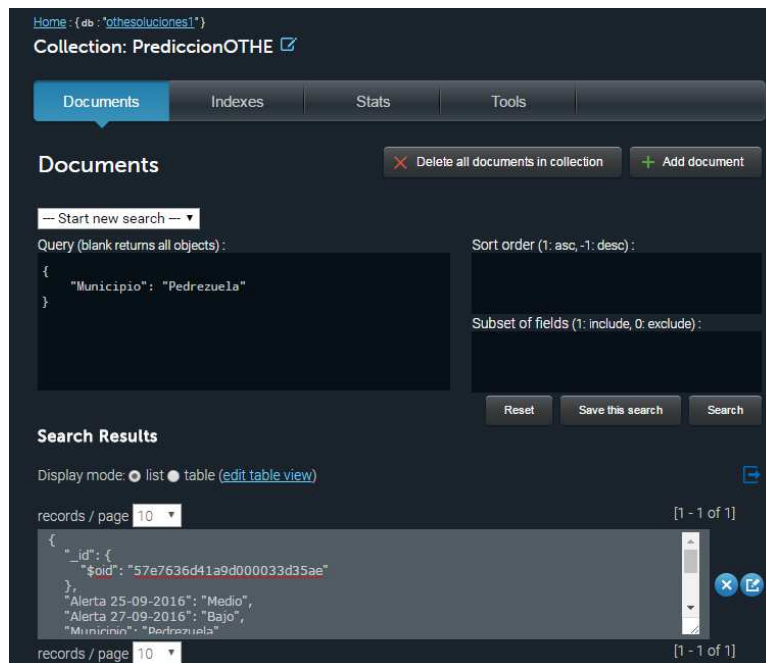
Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Alerta recibida con éxito. Gracias por su colaboración

Municipio	Seleccione un municipio ▼
Nivel	Seleccione un nivel de alerta ▼
Reportar alerta	

Y observamos que la colección ha sido actualizada para ese municipio, asignándole ahora un nivel Medio.

```
{
  "_id": {
    "$oid": "57e7636d41a9d000033d35ae"
  },
  "Alerta 25-09-2016": "Medio",
  "Alerta 27-09-2016": "Bajo",
  "Municipio": "Pedrezuela",
  "Nivel 26-09-2016": 0,
  "Codigo": "28108",
  "Nivel 25-09-2016": 0.9,
  "Alerta 26-09-2016": "Bajo",
  "MunicipioOrig": "Pedrezuela",
  "Nivel 27-09-2016": 0
}
```



Home: (db: 'otbesoluciones')

Collection: PrediccionOTHE

Documents | Indexes | Stats | Tools

Documents ✖ Delete all documents in collection ➕ Add document

Start new search

Query (blank returns all objects):

```
{
  "Municipio": "Pedrezuela"
}
```

Sort order (1: asc, -1: desc):

Subset of fields (1: include, 0: exclude):

Reset Save this search Search

Search Results

Display mode: list table (edit table view)

records / page 10 [1 - 1 of 1]

```
{
  "_id": {
    "$oid": "57e7636d41a9d000033d35ae"
  },
  "Alerta 25-09-2016": "Medio",
  "Alerta 27-09-2016": "Bajo",
  "Municipio": "Pedrezuela",
  "MunicipioOrig": "Pedrezuela",
  "Nivel 25-09-2016": 0.9,
  "Nivel 26-09-2016": 0,
  "Nivel 27-09-2016": 0
}
```

records / page 10 [1 - 1 of 1]

De esta forma se van recalculando los niveles de alerta cada vez que un usuario realiza un reporte en un municipio. Por lo que gracias a la aportación de los usuarios se va dando un mejor servicio ajustándolo más a la realizada del momento.

BÚSQUEDA DE NOTICIAS

En la parte derecha de cada una de las pantallas de la aplicación aparecen las últimas noticias relacionadas con el polen en Madrid durante este año, estas noticias se actualizan diariamente utilizando un motor de búsqueda de google generado a tal efecto que recupera las noticias de la web.

Para poder realizar esto, lo primero que hay que hacer es crear un motor de búsqueda personalizado desde: <https://cse.google.com/cse/> . Para esta aplicación se ha creado un motor llamado “alergias” que se encarga de buscar en los principales periódicos de la zona.



Sitios en los que buscar:

Buscar solo en sitios incluidos

Añadir Eliminar Filtro Etiqueta

1 - 5 de 5

Sitio	Etiqueta
<input type="checkbox"/> http://www.madridiario.es/	
<input type="checkbox"/> http://www.20minutos.es/	
<input type="checkbox"/> http://www.abc.es/	
<input type="checkbox"/> http://elpais.com/	
<input type="checkbox"/> www.elmundo.es	

Últimas Noticias

- ▶ Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- ▶ Más consultas por alergias ante la intensa polinización que sufren ...
- ▶ La capital duplica su récord histórico de concentración de polen de ...
- ▶ Comienza la temporada de alergias impulsada por las altas ...
- ▶ Diez consejos para atenuar los efectos de la alergia al polen - ABC.es
- ▶ Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- ▶ Unos 400.000 madrileños tendrán en estos días fuertes síntomas de ...
- ▶ Los alérgicos al polen pasarán esta primavera peor en Andalucía y ...
- ▶ Sanidad activa su sistema de información diaria sobre los niveles ...
- ▶ El repunte del polen de gramíneas y olivo complica la primavera a ...

Una vez creado el motor de búsqueda, Google le asignará un ID, que será necesario para poder ejecutar dicho motor.



Edición
Gratis, con anuncios

Actualizar a Google Site Search (anuncios opcionales)

Detalles

ID de motor de búsqueda

URL pública

Obtener código

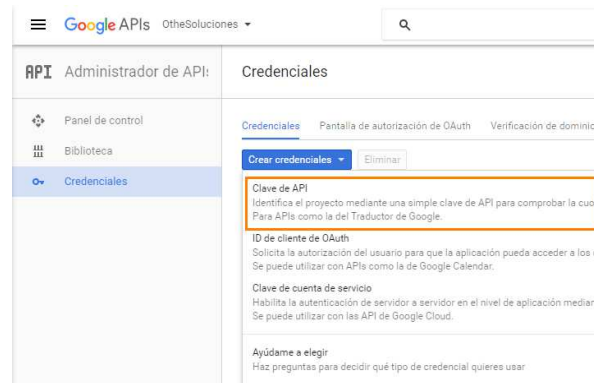
Búsqueda de

ID de tu motor de búsqueda:
006049000923477598507:r2tp5mmc8

Entrada de

si

También es necesario crear una clave para la aplicación de Google (Google API Key), esto se hace a través de la siguiente página: <https://console.developers.google.com> En la opción de menú Credenciales se selecciona el botón Crear Credenciales > Clave de API



De esta forma se asignará una clave:



Ahora ya podremos llamar a la siguiente función utilizando la clave y el id generados y con las palabras clave que deseemos buscar.

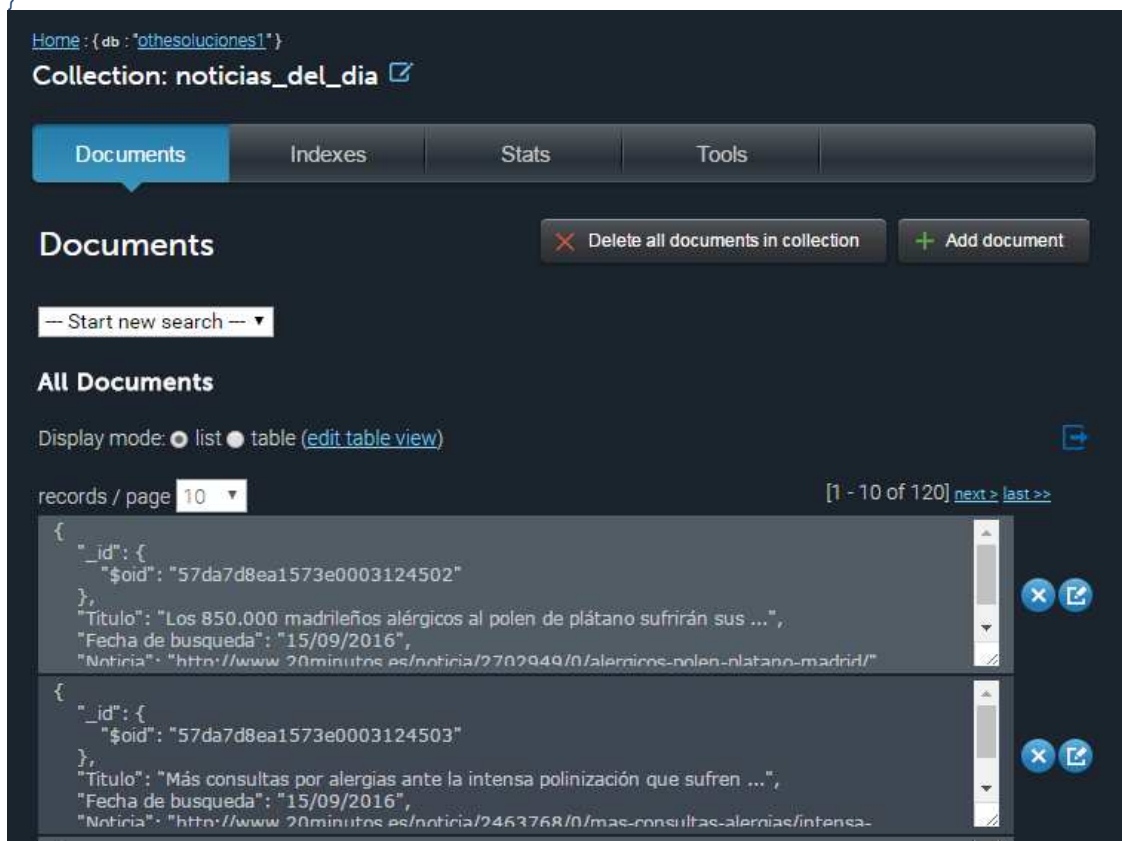
```
my_api_key = "Google API key"
my_cse_id = "Custom Search Engine ID"

def google_search(search_term, api_key, cse_id, **kwargs):
    service = build("customsearch", "v1", developerKey=api_key)
    res = service.cse().list(q=search_term, cx=cse_id, **kwargs).execute()
    return res['items']
```

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Una vez realizado esto almacenamos los resultados obtenidos en la colección de MongoDB **Noticias_del_dia** para el día en el que se realiza la búsqueda almacenando únicamente el título y la url:

```
{
  "_id": {
    "$oid": "57da7d8ea1573e0003124502"
  },
  "Titulo": "Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...",
  "Fecha de búsqueda": "15/09/2016",
  "Noticia": "http://www.20minutos.es/noticia/2702949/0/alergicos-polen-platano-madrid/"
}
```



Home: { db: "othsoluciones1" }

Collection: noticias_del_dia

Documents | Indexes | Stats | Tools

Documents

— Start new search —

All Documents

Display mode: ☒ list ☐ table (edit table view)

records / page 10 [1 - 10 of 120] next > last >>

```
{
  "_id": {
    "$oid": "57da7d8ea1573e0003124502"
  },
  "Titulo": "Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...",
  "Fecha de búsqueda": "15/09/2016",
  "Noticia": "http://www.20minutos.es/noticia/2702949/0/alergicos-polen-platano-madrid/"
}
```

```
{
  "_id": {
    "$oid": "57da7d8ea1573e0003124503"
  },
  "Titulo": "Más consultas por alergias ante la intensa polinización que sufren ...",
  "Fecha de búsqueda": "15/09/2016",
  "Noticia": "http://www.20minutos.es/noticia/2463768/0/mas-consultas-alergias/intensa-"
}
```

APLICACIÓN WEB

Hemos dividido la aplicación web en cuatro capas:

- 1.- Cabecera
- 2.- Barra Lateral
- 3.- Pie de página
- 4.- Contenido de la aplicación



Cabecera:

Consta de 5 pestañas con las opciones de navegación que ofrece la aplicación y de un panel que carga una imagen en función de la pestaña en la que nos encontremos.



Barra Lateral:

Que a su vez hemos dividido en dos apartados distintos: Links de Interés y Últimas Noticias.

En Links de Interés, ofrecemos los enlaces a cuatro webs que consideramos pueden ser de utilidad para los usuarios:

Links de interés

- ▶ Sociedad Española de Alergología (SEAIC)
- ▶ Portal de Salud Comunidad de Madrid
- ▶ LaAlergia.com
- ▶ AlergiaAlPolen.com

En Últimas Noticias, se cargan diez de las noticias encontradas en la red gracias al motor de búsqueda implementado en esta aplicación, que se ejecuta forma diaria:

Últimas Noticias

- ▶ Más consultas por alergias ante la intensa polinización que sufren ...
- ▶ Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- ▶ Sanidad activa su sistema de información diaria sobre los niveles ...
- ▶ Unos 400.000 madrileños tendrán en estos días fuertes síntomas de ...
- ▶ El repunte del polen de gramíneas y olivo complica la primavera a ...
- ▶ Los alérgicos al polen pasarán esta primavera peor en Andalucía y ...
- ▶ Comienza la temporada de alergias impulsada por las altas ...
- ▶ La capital duplica su récord histórico de concentración de polen de ...
- ▶ Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- ▶ Diez consejos para atenuar los efectos de la alergia al polen - ABC.es

Pie de página:



Donde incluidos los links a la validación del código HTML y CSS utilizado en nuestra web, junto con el logo de Othe Soluciones.

Contenido de la aplicación:

Donde ofrecemos la información relacionada con la pestaña seleccionada:

Pestaña HOME:

HOME

NIVELES DEL DÍA

PREDICCIONES

¡REPÓRTANOS!

NOTIFICACIONES



Bienvenidos

Bienvenidos a <https://gramineas-madrid.herokuapp.com/> web dedicada a la detección de brotes alérgicos de gramíneas en la Comunidad de Madrid.

Podrá consultar los datos meteorológicos y de calidad del aire en cada uno de los municipios de nuestra comunidad recogidos hoy mismo, así como la previsión del nivel de alerta.

Además podrá reportarnos alertas sobre brotes que se estén produciendo en cualquiera de estos municipios ayudándonos a mejorar nuestro modelo de predicción.

También podrán recibir en su correo electrónico la predicción de brotes alérgicos para las fechas y municipios que decida.

Para cualquier sugerencia sobre la información que se ofrece en este servicio contactar con:

othsoluciones@gmail.com

Links de interés

- ▶ Sociedad Española de Alergología (SEAIC)
- ▶ Portal de Salud Comunidad de Madrid
- ▶ LaAlergia.com
- ▶ AlergiaAlPolen.com


Últimas Noticias

- ▶ Más consultas por alergias ante la intensa polinización que sufren ...
- ▶ Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- ▶ Sanidad activa su sistema de información diaria sobre los niveles ...

Esta es la página que contiene la presentación de la aplicación y ofrece la información de contacto con los desarrolladores.

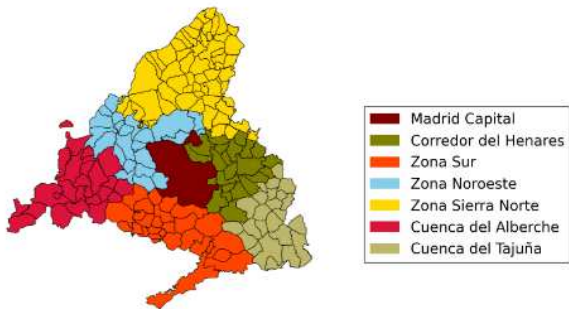
Pestaña Niveles del día:

HOME
NIVELES DEL DÍA
PREDICCIONES
¡REPÓRTANOS!
NOTIFICACIONES



Municipios de la Comunidad de Madrid

Seleccione un municipio de la Comunidad de Madrid para visualizar sus datos meteorológicos y de calidad del aire correspondientes al día: 24-09-2016



Madrid Capital	Madrid	
Corredor del Henares	Zona Noroeste	Cuenca del Alberche
<ul style="list-style-type: none"> Ajalvir Alcalá de Henares Alcobendas Algete Anchuelo Arganda del Rey Camarma de Esteruelas 	<ul style="list-style-type: none"> Alpedrete Boadilla del Monte Cercedilla Collado Mediano Collado Villalba Colmenarejo Colmenar Viejo 	<ul style="list-style-type: none"> Aldea del Fresno Brunete Cadalso de los Vidrios Cenicientos Chapinería Colmenar del Arroyo Fresnedillas de la Oliva

Links de interés

- Sociedad Española de Alergología (SEAIC)
- Portal de Salud Comunidad de Madrid
- LaAlergia.com
- AlergiaAlPolen.com

Últimas Noticias

- Más consultas por alergias ante la intensa polinización que sufren ...
- Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- Comienza la temporada de alergias impulsada por las altas ...
- La capital duplica su récord histórico de concentración de polen de ...
- Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- Diez consejos para atenuar los efectos de la alergia al polen - ABC.es
- Los alérgicos al polen pasarán esta primavera peor en Andalucía v ...

En esta pestaña se muestra un mapa de la Comunidad de Madrid con las siete zonas en las que hemos dividido los municipios en función de su ubicación geográfica: Madrid Capital, Corredor del Henares, Zona Sur, Zona Noroeste, Zona Sierra Norte, Cuenca del Alberche y Cuenca del Tajuña.

Para cada una de estas zonas, mostramos el listado de municipios que la componen, ofreciendo para cada uno de ellos un link a una nueva pantalla:


HOME

NIVELES DEL DÍA



PREDICCIONES

¡REPÓRTANOS!

NOTIFICACIONES



Alcalá de Henares

Datos Meteorológicos	
Temperatura máxima (°C)	28
Temperatura mínima (°C)	15
Humedad relativa máxima (%)	90
Humedad relativa mínima (%)	30
Precipitaciones (%)	0
Velocidad del Viento (km/h)	0

Datos Calidad del Aire	
Dióxido de Nitrógeno (µg/m³)	32.86
Partículas en Suspensión < PM10 (µg/m³)	16.67
Partículas en Suspensión < PM2,5 (µg/m³N)	7.5
Monóxido de Carbono (mg/m³)	0.5
Concentración de Ozono (µg/m³)	30.57
Dióxido de Azufre (µg/m³)	4.0
Monóxido de Nitrogeno (µg/m³)	6.71

Volver

Links de interés

- Sociedad Española de Alergología (SEAIC)
- Portal de Salud Comunidad de Madrid
- LaAlergia.com
- AlergiaAlPolen.com

Últimas Noticias

- Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- Diez consejos para atenuar los efectos de la alergia al polen - ABC.es
- Comienza la temporada de alergias impulsada por las altas ...
- La capital duplica su récord histórico de concentración de polen de ...
- Sanidad activa su sistema de información diaria sobre los niveles ...
- Los alérgicos al polen pasarán esta primavera peor en Andalucía y ...
- Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- Unos 400.000 madrileños tendrán en estos días fuertes síntomas de ...
- Más consultas por alergias ante la intensa polinización que sufren ...


En esta nueva pantalla, de forma ampliada aparece un gráfico con la silueta del municipio y al lado su ubicación resaltada dentro del mapa de la Comunidad de Madrid.

Además se presentan dos tablas con la información meteorológica y de calidad del aire del día en el que nos encontramos, esta información se encuentra almacenada en nuestra base de datos.

Estos mapas, junto con el anterior de la Comunidad de Madrid y los que se muestran en la pantalla de Predicciones por municipio, fueron generados y cargados en la base de datos al ejecutar el fichero **generaMapas.py** previamente, ya que los tiempos de carga eran más rápidos cargándolos desde BBDD que generándolos en ejecución.

Pestaña Predicciones:

HOME
NIVELES DEL DÍA
PREDICCIONES
¡REPÓRTANOS!
NOTIFICACIONES



Predicciones del nivel de alérgenos por municipio

Seleccione un municipio para visualizar la predicción del nivel de brote alérgico de gramíneas para los próximos 3 días.

Madrid	Alto
Villa del Prado	Medio
Alcorcón	Medio
Rascafría	Bajo
Valdemaqueda	Bajo



Madrid	Medio
Rascafría	Bajo
Valdemaqueda	Bajo
Manzanares el Real	Bajo
Meco	Bajo

Rascafría	Bajo
Valdemaqueda	Bajo
Manzanares el Real	Bajo
Meco	Bajo
Mejorada del Campo	Bajo

Links de interés

- Sociedad Española de Alergología (SEAIC)
- Portal de Salud Comunidad de Madrid
- LaAlergia.com
- AlergiaAlPolen.com

Últimas Noticias

- Más consultas por alergias ante la intensa polinización que sufren ...
- Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- Comienza la temporada de alergias impulsada por las altas ...
- La capital duplica su récord histórico de concentración de polen de ...
- Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- Diez consejos para atenuar los efectos de la alergia al polen - ABC.es
- Los alérgicos al polen pasarán esta primavera peor en Andalucía y ...
- Unos 400.000 madrileños tendrán en estos días fuertes síntomas de ...
- El repunte del polen de gramíneas y olivo complica la primavera a ...
- Sanidad activa su sistema de información diaria sobre los niveles ...

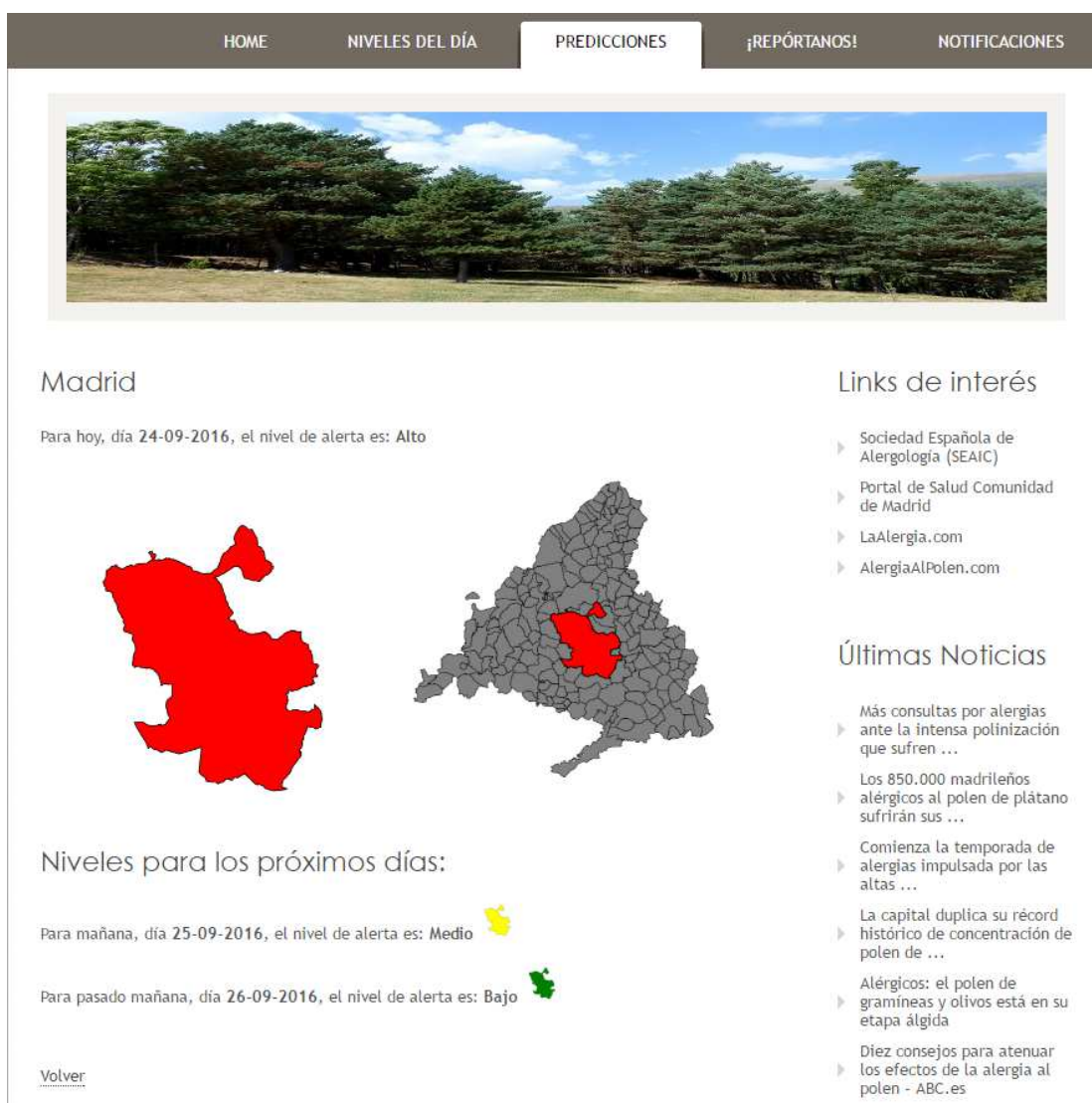
Esta pestaña la podríamos dividir en tres secciones:

La inferior muestra dos tablas con los cinco municipios que tienen un mayor nivel de alerta para los dos días siguientes al actual. Este nivel, ha sido calculado por nuestro algoritmo predictivo.

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

La sección intermedia muestra una tabla con la misma información calculada para el día en el que nos encontramos, junto con un mapa de la Comunidad de Madrid, cuyos municipios están coloreados de forma dinámica en función del nivel de alerta existente en la base de datos.

La sección superior muestra un formulario con un combo que carga el nombre de todos los municipios de la Comunidad, y que al seleccionar uno de ellos y pulsar sobre el botón Ver Predicción nos redirigirá a la siguiente pantalla:



HOME NIVELES DEL DÍA PREDICCIONES ¡REPÓRTANOS! NOTIFICACIONES

Madrid

Para hoy, día 24-09-2016, el nivel de alerta es: Alto

Links de interés

- Sociedad Española de Alergología (SEAI)
- Portal de Salud Comunidad de Madrid
- LaAlergia.com
- AlergiaAlPolen.com

Últimas Noticias

- Más consultas por alergias ante la intensa polinización que sufren ...
- Los 850.000 madrileños alérgicos al polen de plátano sufrirán sus ...
- Comienza la temporada de alergias impulsada por las altas ...
- La capital duplica su récord histórico de concentración de polen de ...
- Alérgicos: el polen de gramíneas y olivos está en su etapa álgida
- Diez consejos para atenuar los efectos de la alergia al polen - ABC.es

Niveles para los próximos días:

Para mañana, día 25-09-2016, el nivel de alerta es: Medio

Para pasado mañana, día 26-09-2016, el nivel de alerta es: Bajo

Volver

En la que se muestra para el municipio seleccionado, su nivel de alerta para el día actual junto con dos mapas, el del propio municipio y el de su ubicación en la Comunidad, cuyo color depende del valor del nivel de alerta calculado para el día actual.

Además también aparece la predicción del nivel de alertas para los dos días siguientes cargando igualmente, pero con un formato más reducido del mapa del municipio y con el color correspondiente a su nivel de alerta.

Pestaña ¡Repórtanos!

HOME	NIVELES DEL DÍA	PREDICCIONES	¡REPÓRTANOS!	NOTIFICACIONES
------	-----------------	--------------	---------------------	----------------



Repórtanos alertas

Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Municipio	<div style="border: 1px solid #ccc; padding: 2px;"> Seleccione un municipio ▼ </div>
Nivel	<div style="border: 1px solid #ccc; padding: 2px;"> Seleccione un nivel de alerta ▼ </div>
<div style="border: 1px solid #ccc; padding: 2px 10px;">Reportar alerta</div>	

No existen alertas notificadas por nuestros usuarios durante el día 24-09-2016

Links de interés

- ▶ Sociedad Española de Alergología (SEAIC)
- ▶ Portal de Salud Comunidad de Madrid
- ▶ LaAlergia.com
- ▶ AlergiaAlPolen.com

Últimas Noticias

▶ Más consultas por alergias ante la intensa polinización que sufren ...

En esta pantalla se muestra un formulario que permite al usuario introducir una alerta para un municipio de la Comunidad de Madrid, junto con una tabla con las alertas reportadas al sistema durante el día (en caso de no haberse reportado ninguna, se muestra un mensaje con esta información).

En caso de realizar una selección correcta en ambos combos se procederá a la actualización de la predicción del nivel de alerta del municipio en nuestro sistema y se mostrará un mensaje de confirmación de alta de la alerta:

Repórtanos alertas

Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Municipio	<div style="border: 1px solid #ccc; padding: 2px;"> Collado Villalba ▼ </div>
Nivel	<div style="border: 1px solid #ccc; padding: 2px;"> Alto ▼ </div>
<div style="border: 1px solid #ccc; padding: 2px 10px;">Reportar alerta</div>	

No existen alertas notificadas por nuestros usuarios durante el día 24-09-2016

Vista de la pestaña ¡Repórtanos! Tras alta correcta:

Repórtanos alertas

Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Alerta recibida con éxito. Gracias por su colaboración

Municipio	Seleccione un municipio ▼
Nivel	Seleccione un nivel de alerta ▼
	Reportar alerta

Alertas notificadas por nuestros usuarios durante el día 24-09-2016:

Municipio	Nivel de Alerta Reportada	Hora
Collado Villalba	Alto	19:56:11

En caso contrario se mostrará un mensaje solicitando al usuario que realice una selección correcta:

Vista de la pestaña ¡Repórtanos! Tras alta incorrecta

Repórtanos alertas

Si quieres ayudarnos a mejorar nuestra predicción, puedes hacerlo si nos reportas una alerta del municipio en el que te encuentres.

Por favor, seleccione un municipio correcto


Por favor seleccione un nivel correcto

Municipio	Seleccione un municipio ▼
Nivel	Seleccione un nivel de alerta ▼
	Reportar alerta

No existen alertas notificadas por nuestros usuarios durante el día 24-09-2016

Pestaña Notificaciones

HOME	NIVELES DEL DÍA	PREDICCIONES	¡REPÓRTANOS!	NOTIFICACIONES
------	-----------------	--------------	--------------	----------------



Recibe notificaciones

Si quieres recibir en tu correo electrónico un email con la predicción del nivel de alertas del municipio que desees, por favor rellena este formulario.

Email	<input type="text"/>
Municipio	<div> <div>Seleccione un municipio</div> <div></div> </div>
Periodicidad	<div> <div>Fecha desde:</div> <div>25/09/2016</div> </div> <div> <div>Fecha hasta:</div> <div>25/09/2016</div> </div>
	<input type="button" value="Realizar suscripción"/>

Links de interés

- ▶ Sociedad Española de Alergología (SEAIC)
- ▶ Portal de Salud Comunidad de Madrid
- ▶ LaAlergia.com
- ▶ AlergiaAlPolen.com

Últimas Noticias

▶ Más consultas por alergias ante la intensa polinización que sufren ...

En esta pantalla se muestra un formulario que permite a un usuario recibir en su correo electrónico un email durante el periodo elegido con el nivel de alertas del municipio seleccionado.

Si alguno de los campos del formulario se rellena de forma incorrecta, se mostrará un mensaje por pantalla solicitando la corrección de dichos campos.

Recibe notificaciones

Si quieres recibir en tu correo electrónico un email con la predicción del nivel de alertas del municipio que desees, por favor rellena este formulario.

Por favor, introduzca una dirección de correo electrónico correcta

Por favor, seleccione un municipio correcto

Email	<input type="text"/>
Municipio	Seleccione un municipio ▼
Periodicidad	Fecha desde: 25/09/2016
	Fecha hasta: 25/09/2016
	<input type="button" value="Realizar suscripción"/>

El campo email validará que el formato introducido del correo electrónico sea correcto.

El campo municipio validará que se seleccione correctamente un municipio del desplegable.

Y para las fechas, se cargan dos componentes datepicker de **jQuery** que siempre van a cumplir las siguientes validaciones: la fecha desde siempre va a ser mayor o igual que la del día siguiente en el que nos encontramos y la fecha hasta siempre va a ser mayor igual que la fecha seleccionada en el campo fecha desde.

Si las tres validaciones se realizan con éxito, se dará de alta la suscripción en nuestro sistema mostrándose por pantalla un mensaje de confirmación.

Así, durante el periodo indicado, enviaremos al usuario un email con la información del nivel de alerta del municipio del día, junto con la de los dos días posteriores y la fecha en la que caducará la suscripción creada.

Además, si un usuario se suscribe durante el mismo periodo a las notificaciones de dos o más municipios, recibirá la información de todos los municipios en un único email.

Recibe notificaciones

Si quieres recibir en tu correo electrónico un email con la predicción del nivel de alertas del municipio que desees, por favor rellena este formulario.

Email	<input type="text" value="lothesoluciones@gmail.com"/>
Municipio	<input type="text" value="Leganés"/>
Periodicidad	Fecha desde: <input type="text" value="25/09/2016"/> Fecha hasta: <input type="text" value="27/09/2016"/>
	<input type="button" value="Realizar suscripción"/>

Recibe notificaciones



Si quieres recibir en tu correo electrónico un email con la predicción del nivel de alertas del municipio que desees, por favor rellena este formulario.





Servicio de notificación recibido con éxito. Gracias por su colaboración


Email	<input type="text"/>
Municipio	<input type="text" value="Seleccione un municipio"/>
Periodicidad	Fecha desde: <input type="text" value="25/09/2016"/> Fecha hasta: <input type="text" value="25/09/2016"/>
	<input type="button" value="Realizar suscripción"/>

SISTEMA DE GENERACIÓN DE ALERTAS DE BROTES ALÉRGICOS POR GRAMÍNEAS

Ejemplo de email recibido el 25-09-2016:

25-09-2016. Servicio de Notificaciones Recibidos x  

 **othesoluciones@gmail.com** 7:45 (hace 6 horas)   

para mí 

Buenos días,

Estas son las notificaciones que ha solicitado:

Moralzarzal:

El Nivel de Alerta de Gramíneas para el día 25-09-2016 es: **Bajo**

El Nivel de Alerta de Gramíneas para el día 26-09-2016 es: **Bajo**

El Nivel de Alerta de Gramíneas para el día 27-09-2016 es: **Bajo**

Recibirá esta notificación hasta el: **29-09-2016**

Leganés:

El Nivel de Alerta de Gramíneas para el día 25-09-2016 es: **Bajo**

El Nivel de Alerta de Gramíneas para el día 26-09-2016 es: **Bajo**


El Nivel de Alerta de Gramíneas para el día 27-09-2016 es: **Bajo**

Recibirá esta notificación hasta el: **27-09-2016**

Deseamos que pase un gran día.

Para más información puede consultar nuestra web: <http://gramineas-madrid.herokuapp.com/>

Reciba un cordial saludo por parte del equipo de Othe Soluciones



TRABAJO FUTUROS

En esta sección del trabajo se van a enumerar las futuras líneas de trabajo propuestas para continuar el desarrollo de esta aplicación web.

Seguridad

Actualmente el sistema no dispone de un mecanismo de autenticación, y confía en el comportamiento correcto del usuario para la solicitud de la suscripción a las notificaciones y el reporte de alertas en los municipios.

Podría implementarse un sistema de autenticación con login y password incluyendo algún mecanismo de verificación como la confirmación a través de un email del alta en la aplicación. También se podría integrar en el sistema el inicio de sesión mediante APIs de redes sociales como Facebook o Twitter.

Esto podría permitirnos mostrar al usuario un histórico con su actividad en la aplicación, pudiendo incluir funcionalidades de consulta como las alertas reportadas y la gestión de la suscripción a las alertas a las que se ha suscrito el usuario.

Aplicación para dispositivos móviles

Actualmente nuestra aplicación es accesible a través del navegador de cualquier dispositivo, pero consideramos que sería interesante aprovechar el sistema existente para crear sobre una ella aplicación para dispositivos móviles en distintos sistemas operativos como Android o iOS.

Además se podría modificar el comportamiento y la forma de recibir las suscripciones de las alertas que podría hacerse mediante notificación en la app.

Geografía

Nuestra aplicación se ha creado para estudiar el nivel polínico de gramíneas para la Comunidad de Madrid, pero pensamos que se podría ampliar para estudiar de forma completa todos los municipios de España.

Creemos esto ya que, actualmente fuentes de datos como la AEMET o la SEAIC disponen de datos para las distintas Comunidades Autónomas, aunque lógicamente tendríamos que encontrar nuevas fuentes para cada una de ellas como por ejemplo el calendario polínico o los niveles de la calidad de aire ya que ahora trabajamos con las de la Comunidad de Madrid.

Otros pólenes

También podríamos mostrar más niveles de alertas de otras variedades de pólenes generados por otras plantas como por ejemplo, las Urticáceas, las Alternarias o las Cupresáceas... cuyos datos ya recogemos para la Comunidad de Madrid aunque en la actualidad no realizamos ningún tratamiento con ellos.

Procesamiento de Lenguaje Natural

Otra mejora que podría implementarse, sería analizar las noticias recuperadas actualmente con técnicas de NLP, no solo para ordenarlas en función de su importancia, sino también para incluir la valoración obtenida del texto como una nueva variable en nuestro modelo de predicción de alertas.

CONCLUSIÓN

Tras finalizar el proyecto, y teniendo en cuenta que ha sido realizado entre los meses de agosto y septiembre, no hemos conseguido recuperar unos datos de nivel de polen en aire (gr/m³) significativos, esto es algo de esperar teniendo en cuenta el calendario polínico, pero a la hora de implementar el algoritmo predictivo y realizar nuestras pruebas, ha sido un inconveniente ya que siempre hemos obtenido unos valores similares que se traducían en un nivel bajo de alerta. Sería muy recomendable continuar probando el sistema durante al menos un ciclo anual completo, para poder cerciorarnos de la validez del modelo. Así como de la recopilación diaria de noticias, ya únicamente durante el final del invierno y la primavera se consideran las alergias al polen como un tema de actualidad, mientras que en el resto del año a penas si aparece en los medios de comunicación.

Otro de los problemas con el que nos hemos topado, ha sido a la hora de la obtención de datos, ya que algunas de las fuentes oficiales encontradas no permitían la descarga automática de datos vía web scrapping, o bien, no se actualizaban de forma periódica. También intentamos ponernos en contacto con la Red PalinoCAM, del portal de salud de la Comunidad de Madrid, para ver si nos podían proporcionar otra forma de acceso a sus informes, pero no obtuvimos respuesta.

Finalmente queríamos resaltar que hemos intentado aplicar en nuestro proyecto las tres V's que definen el Big Data, así tenemos Variedad en los datos, tanto por su origen y forma como por su formato, este aspecto se puede comprobar por ejemplo, en el tratamiento y visualización de mapas obtenidos en formato shapefile, también hemos intentado que nuestro sistema predictivo fuera Veloz, sobre todo en el momento en el cual el usuario nos reporta una alerta y esta se registra en tiempo real en la aplicación y por último querríamos comentar que aunque el Volumen de datos que se ha manejado para realizar esta aplicación no se puede considerar como Big Data, debido a la limitación de 500MB por el uso de la versión gratuita de Mlab, las técnicas utilizadas para la obtención, el tratamiento o el almacenamiento de los datos han sido implementadas usando tecnología Big Data. No tenemos que olvidar que se trataba de una prueba de concepto, y si en un futuro se decidiesen aplicar las mejoras que se han propuesto, así como los datos históricos, el volumen de datos se incrementaría considerablemente, y los métodos utilizados podrían soportar dicha evolución.

BIBLIOGRAFÍA

- [1] <https://es.wikipedia.org/wiki/Python>
- [2] <https://www.python.org/about/>
- [3] <https://pypi.python.org/pypi/numpy/1.10.1>
- [4] <https://es.wikipedia.org/wiki/NumPy>
- [5] <http://www.numpy.org/>
- [6] [https://en.wikipedia.org/wiki/Pandas_\(software\)](https://en.wikipedia.org/wiki/Pandas_(software))
- [7] <https://pypi.python.org/pypi/pandas/0.17.0/>
- [8] <http://pandas.pydata.org/>
- [9] <https://es.wikipedia.org/wiki/Matplotlib>
- [10] <http://matplotlib.org/>
- [11] <https://pypi.python.org/pypi/xmltodict>
- [12] <http://lxml.de/>
- [13] <https://pypi.python.org/pypi/lxml/3.4.4>
- [14] https://es.wikipedia.org/wiki/Beautiful_Soup
- [15] <https://pypi.python.org/pypi/beautifulsoup4/4.4.1>
- [16] <https://pypi.python.org/pypi/google-api-python-client/>
- [17] <https://developers.google.com/api-client-library/python/apis/customsearch/v1>
- [18] <https://github.com/google/google-api-python-client>
- [19] <https://support.google.com/customsearch/answer/2631040>
- [20] <https://api.mongodb.com/python/current/>
- [21] <https://pypi.python.org/pypi/pymongo>
- [22] <https://apscheduler.readthedocs.io/en/latest/>
- [23] <https://pypi.python.org/pypi/APScheduler/3.2.0>
- [24] <https://pypi.python.org/pypi/pyshp>
- [25] <http://bottlepy.org/docs/dev/index.html>
- [26] <https://pypi.python.org/pypi/bottle/0.12.9>
- [27] [https://en.wikipedia.org/wiki/Bottle_\(web_framework\)](https://en.wikipedia.org/wiki/Bottle_(web_framework))
- [28] <https://www.mongodb.com/es>
- [29] <https://es.wikipedia.org/wiki/MongoDB>
- [30] Documentación facilitada en la clase de bases de datos no relacionales.
- [31] <https://www.heroku.com/>
- [32] <https://mlab.com/>
- [33] <https://jquery.com/>
- [34] <https://es.wikipedia.org/wiki/JQuery>
- [35] <http://conociendogithub.readthedocs.io/en/latest/data/introduccion/>
- [36] <https://es.wikipedia.org/wiki/GitHub>
- [37] <https://github.com>
- [38] <http://www.madrid.org/cs/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadname1=Contentdisposition&blobheadname2=cadena&blobheadervalue1=filename%3Dgramineae.pdf&blobheadervalue2=language%3Des%26site%3DPortalSalud&blobkey=id&blobtable=MungoBlobs&blobwhere=1352809127865&ssbinary=true>
- [39] <http://www.ugr.es/~aerobio/gramineas.htm>

- [40] <http://encuentralainspiracion.es/la-alergia-respiratoria/tipos-de-alergenos/alergia-al-polen/calendario-de-polinizacion/>
- [41] *SEAIC (Sociedad Española de Alergología e Inmunología Clínica)* <http://www.polenes.com/>
- [42] <https://filtrosnosk.com/alergia-en-madrid/>
- [43] <http://www.cronica.com.mx/notas/2011/567707.html>
- [44] http://www.elconfidencial.com/alma-corazon-vida/2016-04-26/primavera-alergia-polen-contaminacion-asma-francisco-feo_1190231/
- [45] <http://www.muyminteresante.es/salud/articulo/la-contaminacion-aumenta-las-alergias>
- [46] http://www.larazon.es/historico/3620-la-polucion-dispara-la-alergia-invernal-ULLA_RAZON_358248#.Ttt1yqixB1OgDI
- [47] <http://www.quo.es/salud/la-contaminacion-dispara-las-alergias>
- [48] <http://revistamotor.eu/index.php/de-calle/mecanica/3331-emisiones-contaminantes-en-motores-diesel-y-gasolina>
- [49] <http://noticias.coches.com/consejos/diesel-contaminacion/164473>
- [50] <http://www.eoi.es/blogs/redinnovacionEOI/2015/10/08/los-motores-diesel-y-la-contaminacion/>
- [51] *Calidad del aire* http://gestiona.madrid.org/azul_internet/html/web/DatosRedAccion.icm?ESTADO_MENU=2_1
- [52] *Valores límite calidad del aire* http://gestiona.madrid.org/azul_internet/html/web/2_3.htm?ESTADO_MENU=2_3
- [53] https://es.wikipedia.org/wiki/Mon%C3%B3xido_de_carbono
- [54] *Partículas en suspensión PM10* <http://www.ecologistasenaccion.es/article5686.html>
- [55] *Partículas en suspensión PM2,5* <http://www.ecologistasenaccion.es/article17842.html>
- [56] *Dióxido de Nitrógeno* http://www.lainformacion.com/salud/que-es-el-dioxido-de-nitrogeno-y-por-que-afecta-tanto-a-la-salud-y-el-medio-ambiente_5dpJTep8mP5CpsEdEP2aj1/
- [57] *Dióxido de Nitrógeno* https://es.wikipedia.org/wiki/Di%C3%B3xido_de_nitr%C3%B3geno
- [58] *Monóxido de nitrógeno* <http://quimica.laguia2000.com/quimica-inorganica/contaminacion-por-monoxido-de-nitrogeno>
- [59] *Monóxido de Nitrógeno* https://www.ecured.cu/Mon%C3%B3xido_de_nitr%C3%B3geno
- [60] *Dióxido de Azufre* <http://www.saludgeoambiental.org/dioxido-azufre-so2>
- [61] *Ozono* <http://www.greenfacts.org/es/ozono-o3/>
- [62] *Predicción meteorológica AEMET* <http://www.aemet.es/es/eltiempo/prediccion/municipios?p=28&w=t>
- [63] *Alergias estacionales* <http://kidshealth.org/es/parents/seasonal-allergies-esp.html>
- [64] <http://stackoverflow.com/questions/37083058/programmatically-searching-google-in-python-using-custom-search>
- [65] <http://www.alergiainfantillafe.org/gramineas.htm>

[66] <http://www.alergiafbvva.es>