**Pasos para la creación y ejecución del script**

En primer lugar, se debe tener una versión de Python 3 (preferiblemente la última) instalada en el sistema operativo. Esto lo podemos hacer desde el sitio web oficial de Python:

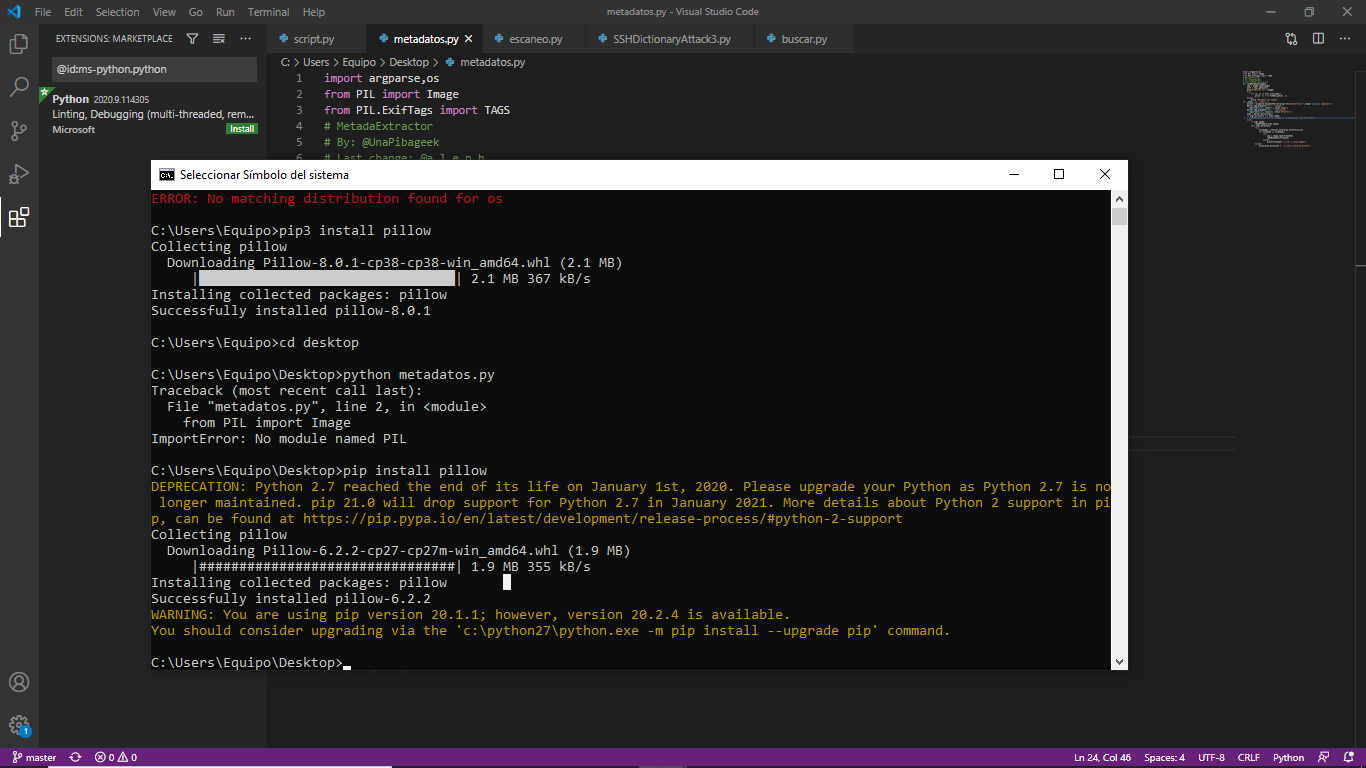
<https://www.python.org>

Además, se debe tener un entorno de desarrollo instalado como lo es Visual Studio Code, o Atom, o Sublime Text 3. Incluso se podría desarrollar desde el block de notas de Windows y al guardar, guardar con la terminación deseada. Para el desarrollo de este proyecto se usó Visual Sudio Code que puede ser descargado desde su página web oficial:

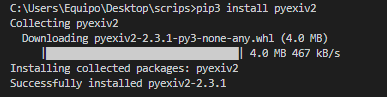
<https://code.visualstudio.com>

Una vez con esto, se abre la terminal, bien sea desde el entorno de desarrollo o directamente desde el sistema operativo y se instalan las librerías Pillow, pyexiv y image\_search. Las dos primeras se usan en la extracción de los metadatos de los archivos de imagen y la última se utiliza en la descarga de imágenes mediante scrapping de Bing. Es importante resaltar que funcionaría también con google, si no fuese porque recientemente la compañía cambió la estructura del source code de la búsqueda. Los comandos para instalar las librerías son

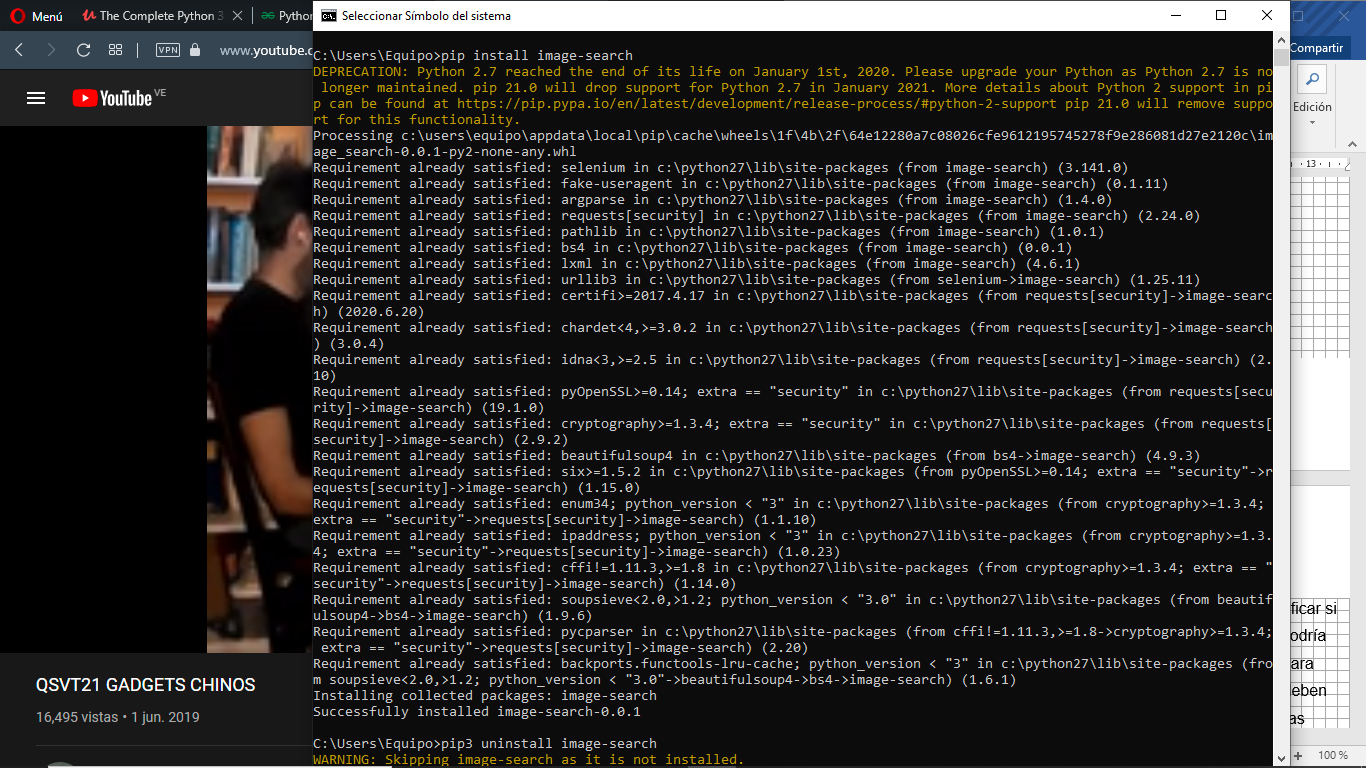
pip install Pillow



pip install pyexiv2



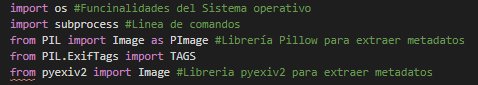
pip install image\_search

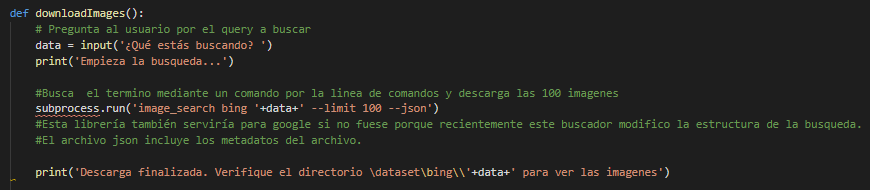


Respecto a la instalación de las librerías, es importante verificar si se tiene instalado o no previamente python2.7, pues si es así se podría tener problemas de compatibilidad respecto a la herramienta pip para instalar librerías. Dado el caso que así sea, ambas herramientas deben tener direcciones distintas en el PATH. Es por esta razón que en las capturas se evidencia que se usa en realidad pip3 y no pip. Dado el caso en que solo se encuentre instalado la última versión de Python 3, esto no es necesario.

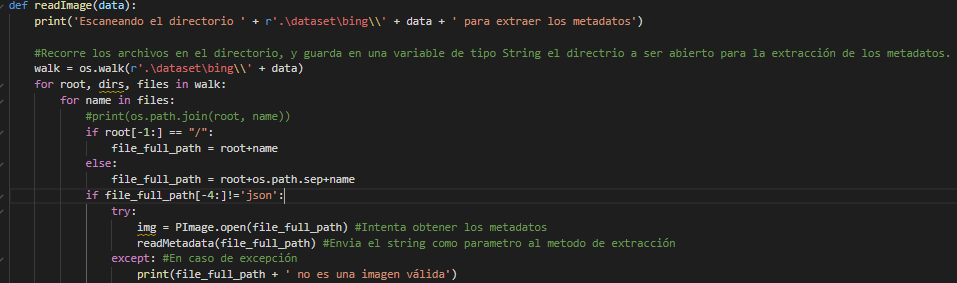
En el particular de image-search, esta librería es instalada como un programa al que se accede mediante la consola para hacer la búsqueda y descargar las imágenes mediante argumentos. Puede ser instalado tanto en python2 como en python3.

Para la creación del Script, en primer lugar se importan las librerías necesarias. Luego se crea el primer método, llamado downloadImages() en el cual se le pregunta al usuario por input “¿Qué estás buscando?” y se guarda la respuesta en una variable de tipo String. Luego se llama a la terminal y se le envía el comando para buscar, concatenado con el termino a buscar. También recibe como parámetro la cantidad de imágenes a descargar y el archivo json de metadatos. Este método retorna el término buscado.

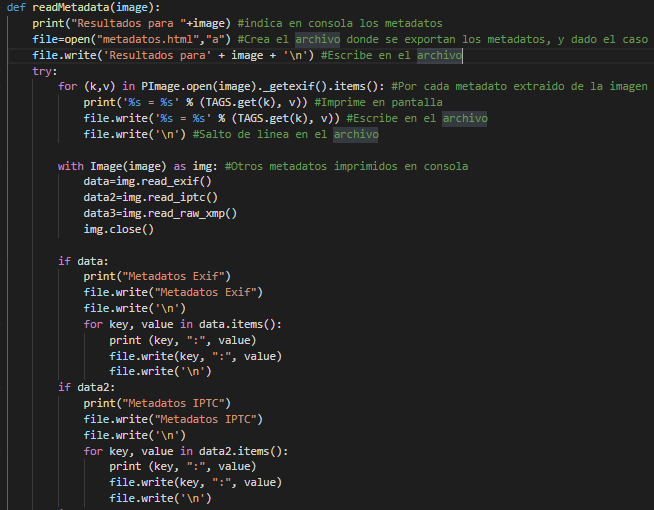


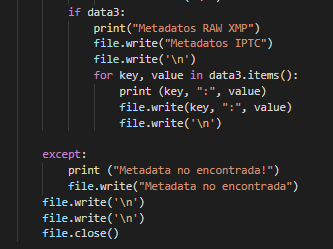


Después se encuentra el método readImage(data), el cual recibe el término como parámetro. Este recorre el directorio donde se encuentran las imágenes, y construye las direcciones de cada uno de los archivos para intentar abrir cada uno de estos con el extractor de metadatos, y luego envía la dirección al siguiente método, llamado readMetadata(image). En caso de excepción se indica por consola que el archivo no es una imagen válida.

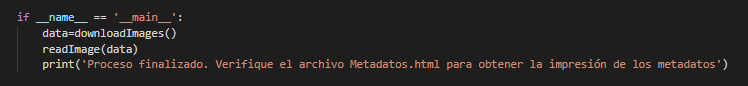


Por consiguiente, se encuentra el método readMetadata(image), el cuál recibe por parámetro la dirección de la imagen. Este método crea el archivo “metadatos.html” y dado el caso de que el archivo exista, lo abre con el puntero al final del archivo, de manera que pueda ser sobre escrito. Una vez allí, intenta abrir nuevamente el archivo con la clase Image de la librería Pillow, e itera sobre el diccionario que el método devuelve, para ir imprimiendo en pantalla y en el archivo html los tags y el valor correspondiente. Luego hace lo mismo con los metadatos de la librería pyexiv2. Al final, cierra el archivo creado.





Finalmente se encuentra el main que guarda la variable data el valor retornada por downloadImages(), la cual pasa al método readImage() y una vez finalizado el proceso, lo indica en pantalla.

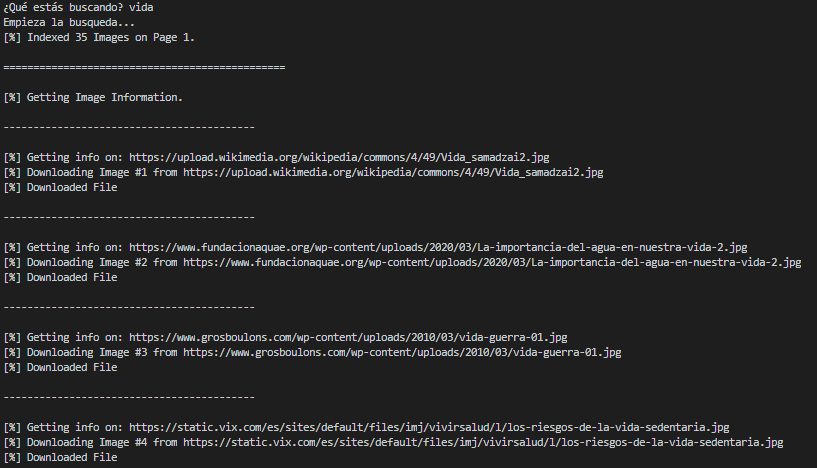


Para la ejecución del archivo, se deben tener instaladas las librerías anteriormente descargadas, y se llama por consola el comando en la raíz del archivo

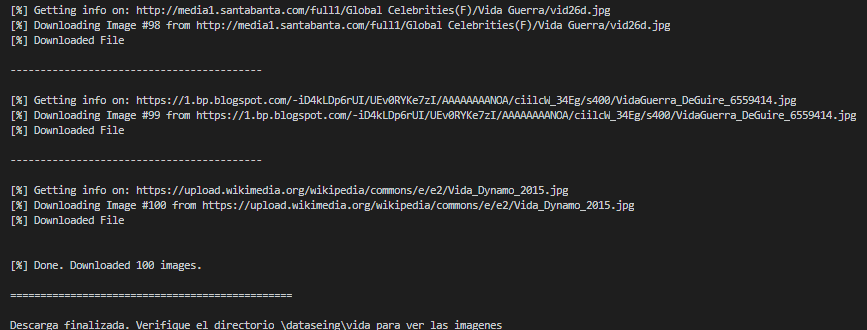
python metadatos3.py



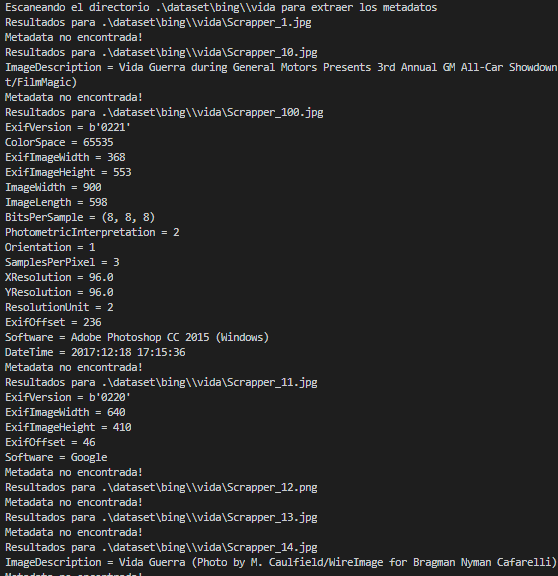
y a partir de allí se ejecuta el programa. El programa solicita el término y empieza la búsqueda y descarga de imágenes



…



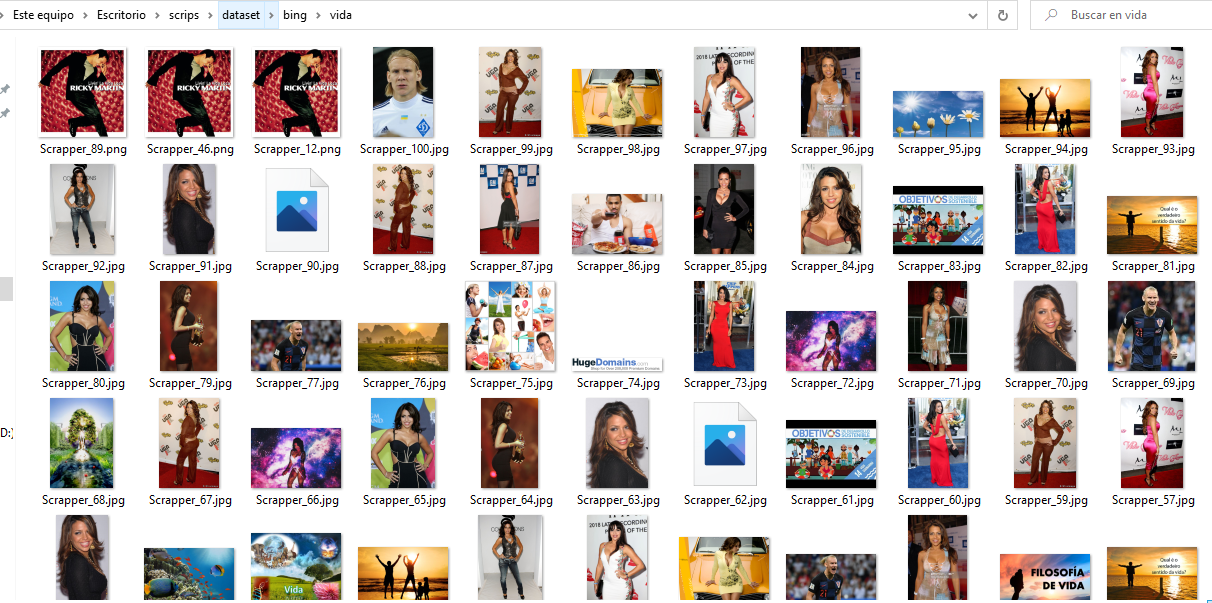
Una vez finalizada la descarga, empieza el escaneo y extracción de metadatos por cada archivo.



Luego finaliza el proceso y se indica esto por consola.



Se verifica la carpeta ./dataset/bing/data y el archivo metadatos.html



Referencias:

<https://comunidad.dragonjar.info/discussion/9528/extraer-metadatos-de-una-imagen-con-python>

<https://es.stackoverflow.com/questions/29405/insertar-salto-de-línea-escribiendo-en-un-txt-python>

<https://stackoverflow.com/questions/4764932/in-python-how-do-i-read-the-exif-data-for-an-image>

<https://stackoverflow.com/questions/11331778/find-backslash-in-a-string-python>

<https://stackoverflow.com/questions/63194917/python-how-to-solve-oserror-errno-22-invalid-argument>

<https://libraries.io/pypi/image-search>

<https://pypi.org/project/image-search/#description>

<https://github.com/LeoHsiao1/pyexiv2/blob/master/docs/Tutorial.md>

<https://www.tutorialspoint.com/python/os_walk.htm>

<https://uniwebsidad.com/libros/algoritmos-python/capitulo-9/utilizando-diccionarios-en-python>