**Inspeccionando los repositorios**

- Planificación del trabajo

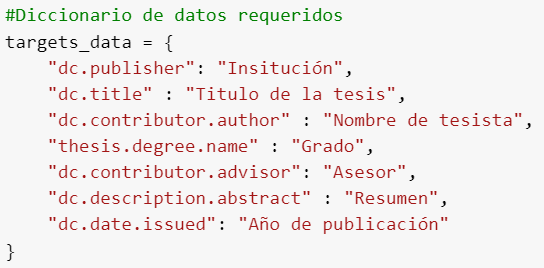
- Inspeccionar los repositorios de la UNALM, PUCP, UPC y U. LIMA de la carrera de Estadística Informática y afines.

- Accedemos a la metadata de las tesis de los repositorios.

- La información que se busca extraer de cada tesis es: Institución, título de la tesis, nombre de tesista, grado, asesor, resumen y año.

**Automatización de la descarga de la información de la tesis**

- Se crea diccionario de la metadata con los datos requeridos.



- Unificar el scraper de las 4 universidades.

* Automatización de la búsqueda en UNALM.
* Automatización de la búsqueda en PUCP.
* Automatización de la búsqueda en UPC.
* Automatización de la búsqueda en ULima.

- Ejecutando todo el código junto.

- Impresión en diccionario de un repositorio solicitado.

* Institución.
* Título de tesis.
* Nombre de tesista.
* Grado.
* Nombre del asesor.
* Resumen.
* Año.

**Estructura de la información descargada.**

* Se crea un array donde se ingresará todas las tesis (cada tesis es un diccionario).
* Recorreremos todas las universidades y recogeremos sus tesis respectivas
* Con el uso de pandas se convertirá en un dataframe.
* Con el df.to\_csv se escribirá un documento en csv.