‘ Cultural heritage data can be complex to explore, every aspect of an objects' history can be a world in itself to discover, which can make it hard to find a starting point when first looking into the data (with or without an objective in mind)’ (Victoria and Albert Museum, 2016)

This portfolio entry was derived from the week 2 lecture on *Getting and cleaning data from API’s*. in this lecture we were introduced to the concept of data wrangling which consisted of merging and reformatting data.

During this lecture we recapped on tabular data and how it is stored (csv/tsv) We were also taught about methods of acquiring data such as API requests and web scraping as well dictionaries and dictionaries formats such as JSON which will come in very handy during this portfolio exercise. The lab exercise was focused on the former and it would be on this topic that this portfolio entry will be based on.

API requests which stands for Application Processing requests is a function where two programmes communicate with each other.

A more detailed way of describing is;

An API, or application programming interface, is a set of rules or protocols that let software applications communicate with each other to exchange data, features and functionality (<https://www.ibm.com/topics/api>, IBM)

The lab exercise consisted of a task to request data from the Metropolitan Museum of Art collection API. We conducted this task using specifically Rest API which is a online service that can be used to access, receive and send data. The aim of this exercise was to explore how an API can prove to be very handy when attempting to focus data acquisition on a small set of data rather than an entire dataset. We used HTTP requests by pointing to a specific URL and proving information getting or sending data and specific queries.

`url` + ? + `keyword1` = `value` & `keyword2` = `value`

Below are a few examples of the results from the Metropolitan museum collection API requests that I made using the following parameters.

api\_url = "https://collectionapi.metmuseum.org/public/collection/v1"

endpoint = "/search"

url = api\_url+endpoint

params = {'q': 'war',

'hasImages': True,

'departmentID': 11,

}

A painting of people riding horses

Description automatically generated

Prior to getting this image there were a number of crucial steps that was taken. As the lecture explained 60% of data science is spent on cleaning data there were a number of steps that were taken before achieving these images. These steps included filtering out columns, putting the data into a dataframe as well as data cleaning.

For this portfolio entry I will perform these same tasks on a collection of data using an API from the Victoria and Albert museum. I will conduct the same processes with the aim of gathering results based on my search queries, organising the results into a dataframe and cleaning and displaying the data.

For this task I decided to choose a search term that which should yield an abundance of results. During a trip to the V&A in January I discovered a number of ready to wear outfits by the famed designer Christian Dior so my queries will be based on search times related to the designer.

The first task I conducted was simply performing a get request on