

## PROGDB'20

### TD 1 : Handling structured files and SQL databases

During this exercise you will work with two databases :

1. `tmdb.db`, a relational database in SQLite format obtained from [tmdb.org](https://www.tmdb.org)
2. `IMDB_dumps`, a database dump in the form of a collection of tab separated value (tsv) files, available on the IMDB web site (c.f., <https://www.imdb.com/interfaces>)

Both databases store information about movies. Your goal is to find the number of movies that the both databases have in common. For the purposes of this exercise, you can identify the movie with its title (`primaryTitle` in the file `title.basics.tsv`) and compare it in a manner that is case insensitive. More precisely, the two titles "A walk in Paris" and "A Walk In Paris" are to be considered equal and representing the same movie. The main objective is to compute correct results efficiently.

You are to submit a Python program `common.py` that outputs the number of movies in common (and no other string). Your program should assume that the files `tmdb.db` and `title.basics.tsv` can be found in the directory `/tmp`. If necessary, your program may create additional file in the current directory (where it is executed).