

## OfficeHub programmerings test

---

### Generelle setup

Programmet er skrevet i Python, jeg har skrevet det med så få ekstra moduler som muligt så du ikke behøver at installere alt muligt. Det eneste kan dog være regex modulet jeg har brugt. Når scriptet er blevet kørt bliver outputtet lagt i en af de tilhørende folders. Der ligger også en fil med testene til den funktion, hvor flere kan blive sat ind.

### How to run

For at køre scripet skal kommandoen køres:

**python3 functions.py**

### Særlige overvejelser

Generelt kunne simple mynstre findes rimelig let med regex og resultatet hentes. Den eneste undtagelse var den sidste funktion om kontakt information, da navner kan have uendelig mange forskellige kombinationer og i eksemplerne også har et stort overlap med titel. Det kunne løses bedre ved brug af NLP funktioner i NLTK biblioteket. Det kunne fx. være en Named entity recognition model.

- Rasmus winther