PDF downloader

*Din kunde har et gammelt Python script, som efter et par år i drift ikke fungerer mere.*

*Du modtager følgende beskrivelse fra din kunde:*

”Kære Konsulent

Som aftalt har du hermed listen (inkl. metadata) ang. de rapporter vi gerne vil have downloaded (”GRI\_2017\_2020”).

Det er adressen i kolonne AL (Pdf\_URL) vi har forsøgt at downloade. Men som jeg forklarede ville det være fedt, at hvis dette link ikke virker, så prøvede programmet linket i kolonne AM. Hvis første link virker, behøver den ikke prøve AM.

Jeg har også vedhæftet vores Python-program, som vi har brugt til at download med. Som sagt kører koden, men den er meget ustabil, og det går til tider meget langsomt.

Desuden er vedhæftet excel-fil med rapporter fra 2006-2016 (”Metadata2006-2016”), hvor man i kolonne AT kan se, hvorvidt en rapport er blevet downloadet eller ej (som vi altså også gerne vil have information om med de nye rapporter). VI behøver IKKE en specificering af, hvorfor rapporten ikke er hentet, blot en variabel der fx antager værdien ”Downloadet” og ”Ikke downloadet”.

Kravspecifikationen for dette projekt er således, at vi ønsker at I laver et program, der effektivt kan downloade alle de rapporter, der har et virkende link, download’er disse pdf-rapporter for os, og deler dem med os via NAS (eller anden måde). De downloadede rapporter skal navngives efter kolonnen BRNummer. Derudover giver I os en liste over hvilke rapporter (fra GRI\_2017\_2020), der er blevet downloadet og hvilke, der ikke er. Den endelige kode afleveres ligeledes til os, da der formegentligt kommer flere rapporter, der skal hentes senere.

Bliver koden lavet i Python eller er det ikke det rigtige program til denne type opgave?”

*Din opgave er nu i dag, at:*

* *Du skal beslutte dig for om du vil opdatere det skrevne Python script eller lavet et nyt. Du har frit valg mht. programmeringssprog, da kunden netop efterspørger din hjælp.*
* *En featureliste/Kravspec skal udarbejdes og godkendes af opgavestilleren inden du går i gang med din løsning.*

*OBS: Din endelige løsning skal inkludere en Readme guide således at enhver studentermedhjælper kan benytte din løsning.*