Kravspecifikation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Navn | Andreas Højgaard-Gammeltoft | | | |
| Opgave | Udvikling af script, der pålideligt og effektivt downloader pdf-filer, angivet i excel-fil. Der skal håndteres alternativ url, hvis den primære ikke virker, og download-status skal registreres. | | | |
| Krav/prioriteter: *Her kan du oplyse de vigtigste elementer i dit software* | Pålidelighed/ordentlig fejlhåndtering, Effektivitet (multithreading), Statusregistrering (”Downloadet” / ”ikke downloadet”), Dokumentation | | | |
| Kodesprog | Python | | | |
| Evt. frameworks | Ingen frameworks. Brug af libraries til download af filer (requests) og læsning af excel-filer (pandas) | | | |
| Estimeret tid | 4 arbejdsdage (5/11 2024 – 8/11 2024) | | | |
|  | | | | |
| Featurelist *Her kan du tilføje hvilket features som skal kodes til dit software* | 1. Download af PDF-filer angivet i GRI\_2017\_2020.xlsx 2. Download af PDF-fil fra alternativ URL ved fejl på primær URL 3. Registrering af downloadstatus til Metadata2006\_2016.xlsx 4. Navngivning af filer efter BRNummer-kolonne 5. Udarbejdning af README-fil med dokumentation | | | |
| Work breakdown *Her kan du bryde dit software ned i arbejdsopgaver og estimere tiden det vil tage at bygge hvert et element* | Arbejdsopgave | Start dato | Forventet slutdato | Tidsforbrug |
| Analyse af opgaven  Download- og fejlhåndtering  Implemenering af multithreading  Statusregistrering  Dokumentation | 5/11 2024  5/11 2024  6/11 2024  7/11 2024  8/11 2024 | 5/112024  6/11 2024  7/11 2024  7/11 2024  8/11 2024 | ½ dag  1 dag  1 dag  ½ dag  1 dag |