**Project: Excel Search Function**

1. **Het bedrijf**

Integrated DNA technologies synthetiseert korte tot lange oligonucleotiden, of anders gezegd stukjes enkelstrengig DNA. Klanten (universiteiten, onderzoekscentra, farma-industrie,…) over de hele wereld kunnen zelf hun eigen oligonucleotide samenstellen. Deze oligo’s worden dan gebruikt in allerlei biomoleculaire applicaties en onderzoeken.

1. **Belangrijke informatie**

***Chemische stoffen***

Bij het synthetiseren van oligo’s moeten veel chemische stoffen worden afgewogen. Denk maar aan modificaties voor op de strengen, stoffen om de streng te doen op groeien,… Al deze stoffen worden gestockeerd in diepvriezers doorheen heel het bedrijf. Op elke stof (zit in een potje) staat een itemnummer en batchnummer vermeld. Het itemnummer is uniek voor die bepaalde stof, terwijl je voor elk itemnummer (elke stof) verschillende batchnummers hebt.



Stof => NaOH

Itemnummer => 1055928 (uniek voor NaoH)

Batchnummer => 0000543698 (uniek voor deze pot NaOH)

Stof => NaOH

Itemnummer => 1055928 (uniek voor NaoH)

Batchnummer => 0000543568 (uniek voor deze pot NaOH)

***SAP***

SAP is een systeem die de hoeveelheid stock van elke stof bijhoudt. Het is dan ook heel belangrijk dat de gegevens in het systeem overeenkomen met de werkelijke stock in de diepvriezers. Op deze manier kan er op tijd stoffen worden bijbesteld als deze dreigen op te raken. Het correct uitschrijven van stoffen na het afwegen is dus uiterst belangrijk.

***Excel Protocol***

Bij elke synthese wordt een apart excel protocol opgesteld. Dit is een voorgemaakte template waarin alle gegevens noodzakelijk voor de synthese worden opgeslagen. Bij het afwegen moet de laborant in dit protocol het itemnummer, batchnummer en hoeveelheid afgewogen stof ingeven.

1. **Het probleem**

***Schetsing probleem***

Wanneer een laborant een stof afweegt moet hij telkens twee dingen doen. De stof uitschrijven in SAP en de hoeveelheid afgewogen stof ingeven in het Excel protocol. Het uitschrijven in SAP wordt af en toe wel eens vergeten waardoor er een gat ontstaat in de gegevens in SAP en de werkelijke stock. Driemaandelijks wordt er een stockcontrole gehouden waarbij de werkelijke stock geteld wordt en deze vergeleken wordt met de SAP-gegevens. Het komt vaak voor dat de werkelijke stock veel lager is dan de gegevens in SAP. Omdat we werken met veel dure stoffen (soms 1000’en euros per gram) moeten deze grote verschillen telkens verantwoord worden aan het management. Hierbij moet men gaan controleren dat alles van die bepaalde stof correct is uitgeschreven (de meest voorkomende fout). De enige manier om dit goed te doen is alle excel-protocols waar die stof in voorkwam te gaan opzoeken en te gaan kijken of dit wel correct is uitgeschreven. Het probleem is dat het enorm veel tijd vergt om te vinden in welke protocols een specifieke stof is gebruikt. Dit probleem heb ik ook zelf eens voorgehad. Er was een verschil van 80 gram van een bepaalde stof en het was dan ook een hele karwei om, via allerlei bronnen, te achterhalen in welke protocols deze stof voorkwam.

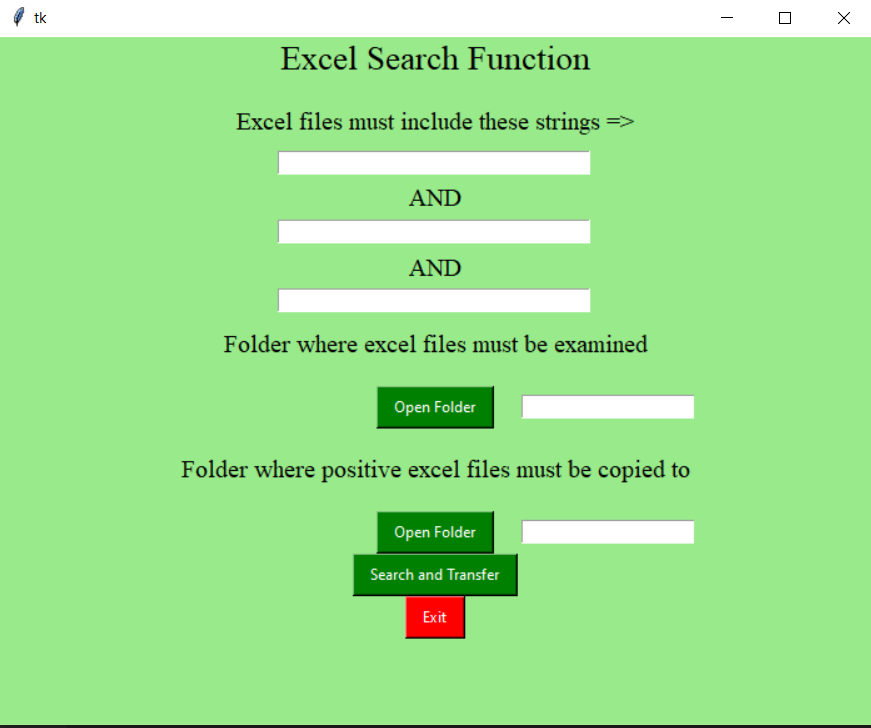
***Oplossing probleem***

Om direct te kunnen weten in welke protocols een bepaalde stof voorkomt, dacht ik eraan een excel zoekfunctie te schrijven in Python. De bedoeling is dat je de naam van je stof kan ingeven en dat het programma een folder met excel files gaat doorzoeken en de excel files die het ingegeven woord bevatten kopieert naar een aparte folder. Op deze manier weet de controleur direct welke protocols deze stof bevatten.

1. **De zoekfunctie**

Hieronder zien we een voorbeeld van de UI van het programma. Het is belangrijk dat de gebruiker alles goed invoert =>

* De gebruiker moet minimum één woord invoeren in het bovenste invoervakje.
  + Twee of drie woorden (van boven naar onder invoeren) is optioneel
    - Functie zoekt excel-files die beide of alledrie de woorden bevat.
    - Het combineren van woorden kan in sommige gevallen handig zijn.
  + Het programma zoekt in de volledige excel-file, niet enkel in de eerste sheet.
* De gebruiker moet een directory selecteren of typen in het invoervakje waarin het programma excel files gaat zoeken.
  + Het programma gaat ook op zoek in alle subdirectory’s van de geselecteerde directory.
* De gebruiker moet een directory selecteren naar waar de positieve excel files worden gekopieerd.

Wanneer alles correct is ingevuld doet het programma zijn werk en geeft een melding met hoeveel positieve excel files zijn gekopieerd naar de gekozen destination folder.