Hoi ik, dit is het bestand waarbij ik ga bijhouden wat ik op de dag heb meegemaakt. Dit kan me later dan helpen bij het maken van de presentaties bij zowel het bedrijfsbezoek als toekomstige stageverslagen. Dit is vooral klad geschreven en veel gepraat, beter te veel dan te weinig denk ik. En ik vergeet snel dingen dus ik kan hier op terugblikken voor enige reminders aan dingen

**Maandag 2 September:**

Een redelijk rustige eerste dag. Ik heb een aantal mensen van de afdeling ontmoet en een rondleiding gehad van het gebouw. Hier om op te merken op de verdieping boven waar ik zit is de IT afdeling. Deze afdeling gaat over de behering van de website en zal ik wanneer het zo ver is langs moeten gaan om te vragen hoe en wat met het zetten van informatie op de website. Hierbij raad ik mezelf aan al voldoende scraping te hebben gedaan zodat ik alvast een voorbeeld kan laten zien aan de mensen daar wat en in welke vorm ik informatie van de websites afhaal. Hierbij zal ik ze dan een beeld geven over hoe de data eruitziet en kan ik feedback vragen over hoe zij met deze data werken. Ook belangrijk om te vragen of ik tijdens dit project de informatie zelf op de website moet zetten, of dat ik het tussenpersoon word die de informatie bij elkaar raapt en doorstuurt naar de IT afdeling.

Hiernaast heb ik nog een vrij uitgebreide uitleg gehad van Jackelien over de doelen van het bedrijf, waarom ze dit doen en waar de problemen liggen. Deze informatie zal ook belangrijk zijn om op te nemen voor het maken van het “domein onderzoek”, deze zal niet verplicht zijn dus ik hoef hem niet in een formaat te schrijven dat school leuk vindt dus vooral hierin verwerken wat MIJ gaat helpen over het stage traject.

**Dinsdag 3 September:**

Tijdens deze dag heb ik mij gefocust op het verzamelen van bronnen en vooral ook vragen stellen over wat voor soort bronnen zij hier vooral op zijn geïnteresseerd, hieruit blijkt dat we dus voornamelijk geïnteresseerd zijn in de bronnen die te maken hebben met zorg- welzijn initiatieven. Wat word hiermee bedoeld is dus voornamelijk activiteiten waarbij burgers zich kunnen aansluiten voor positieve gezondheid. Denk hieraan bij sport activiteiten maar ook sociale bijeenkomsten (ik roep maar wat). Kijk dus aan het begin naar het bij elkaar schrapen van activiteiten rondom Zoetermeer en de zorg- welzijn initiatieven en laten we daarna uitbreiden naar meer. Jackelien en ik hebben besloten dat we eerst willen beginnen met een aantal kern bronnen laten we zeggen 5-10 bronnen die we ECHT willen hebben op de website en vanaf daar uitbreiden. Dit is een bewuste keuze sinds we nog niet de schaal van het project doorhebben, je kan niet van te voren weten hoelang het gaat duren om überhaupt 1 bron volledig werkend te scrapen dus het leek ons beter om minder grootschalig te beginnen en vanaf hier te gaan uitbreiden

**Woensdag 4 september**

Deze dag ben ik niet aanwezig geweest op de werkplek. Dit komt doordat mijn laptop oplader stuk is gegaan en ik dus niet een werkende laptop mee kon nemen. En omdat ik een gesprek met Mirjam had geplant voor het verbeteren van het stageplan. Hierin hebben we wat dingen duidelijk gemaakt zoals het dynamiek tussen ons en wat ik van Mirjam moet verwachten maar ook wat Mirjam van mij verwacht. Dit was een erg fijn gesprek en zorgde voor mij in iedergeval voor een hoop duidelijkheid, note to self KIJK OP BRIGHTSPACE. Hou het in de gaten en kijk goed naar de inlever datums, er zijn 3 stageverslagen die moeten worden ingeleverd TEMPLATE STAAT OP BRIGHTSPACE en voor 4 oktober moet er een bedrijfsbezoekje worden ingeplanned (dit moet je zelf doen). Houd dit “dagboek” dagelijks bij en noteer wat je hebt beleefd en waar je vastloopt. Toekomst ik je gaat er geen spijt van krijgen tijdens het schrijven van de verslagen. Laptop oplader gefikst dat is Nice. Ook wil ik een Excel sheet aanmaken waarin ik bijhoud hoeveel uren ik heb gedraaid op locatie, dit is niet voor school maar zorgt ervoor dat ik duidelijk heb hoe hard ik oprecht bezig ben aan de opdracht. Hierdoor kan ik mijzelf een beetje accountable houden en als ik zie dat ik afrem met progressie kan ik zelf bijsturen

**Donderdag 5 September:**

Ik heb vandaag kennismaking gehad met de communicatie dames van de afdeling. Dit was een kleine 30 minuten om elkaar een beetje te leren kennen en te begrijpen wat zowel ik als hun bijdragen. Rohini doet veel met de sociale media en afbeeldingen, hierbij hebben we al kort besproken dat wanneer ik begin met scrapen ik niet weet of ik daarbij ook afbeeldingen etc. mee kan sleuren. Hierbij zij Rohini dat ze hierbij misschien wel kon bijdragen om afbeeldingen bij te voegen bij de informatie die ik verzamel dus houd die in je achterhoofd. Hiernaast ga ik een takenlijst samenstellen van de taken (bite-sized het liefst) die ik ga moeten uitvoeren en deze doorsturen naar Mirjam, hierbij kunnen we samen dan nadenken of ik het een en ander heb gemist en kan ik daarna een planning voor mezelf gaan maken. Ook moet ik nog een aantal kleine aanpassingen maken aan het stageplan en deze weer opsturen naar Mirjam.

Ook heb ik voor het eerst bij een teambespreking gezeten, hierbij hebben ze besproken over wat ze denken waar ze kunnen verbeteren bij georganiseerde projecten, hier hadden ze bijvoorbeeld het idee dat sommige mensen van organisaties hun meningen vrij anders liggen dan bij de burger (ze hebben een soort “afdeling” waarin één burger kan spreken over verschillende zaken) dit zorgt voor een beeld op dingen en kan waardevolle inzichten opleveren. Maar ook praten ze over simpelere dingen als hoe ze hun powerpoint slides kunnen verbeteren/aantrekkelijker maken voor de kijkers. Vandaag was een ook een soort “gast spreker” zijn naam is Hans en hij heeft MS, hierbij heeft hij het gehad over hoe het leven is met MS wat hij wel en niet kan doen. Maar ook hoe hij zich voelt bij de situatie en hoe hij ervoor zorgt dat hij met een positieve uitkijk en gezonde levenstijl, zelfs met zijn beperkingen nog super positief in het leven staat. Zelf is die nu bezig met het schrijven van een boek om andere in vergelijkbare situaties te kunnen helpen.

**Vrijdag 6 september:**

Een vrij rustige dag, ik heb het stageplan “weer” opgestuurd naar mirjam voor goedkeuring en een begin takenlijst gemaakt met de taken waar ik op dit moment aan kan denken heb geupload. Als ik hem wat meer volledig heb ga ik er denk ik een trello van maken of iets wat overzichtelijker schema van taken wat fijner werkt. Daarnaast heb ik mijzelf weer wat verdiept in de sector en eigenlijk nogsteeds bezig met een bronnen overzicht over welke we willen gebruiken en welke niet. Hierbij moet ik ook jackelien vragen voor een kernlijst van 5-10 bronnen die zij 100% zeker op de website wilt hebben. Het gaat om de informatie die het bedrijf wilt meegeven voor preventieve gezondheid niet de informatie die ik “denk” erop te willen. Redelijk onproductieve dag in de grand scheme of things maar als het beginnetje aan informatie etc. is gelegd en ik kan echt beginnen aan het scrapen/verzamelen van de goeie informatie denk ik dat alles op zijn plek komt en de vooruitgang een stuk sneller gaat.

**Maandag 9 September:**

OV staking heeft ervoor gezorgd dat ik vrij laat pas aankwam op locatie (rond half 12) en ik heb het grootste gedeelte van mijn ochtend dus stilgezeten op een bankje hopend dat de metro eindelijk komt opdagen. Daarnaast heb ik nog wel iets nuttigs gedaan vandaag. We hebben een website gevonden die we dus 100% willen gebruiken, Zoetermeerwijzer.nl deze ben ik alvast gaan bestuderen op informatie die we willen extraheren van de website en ik heb een soort omgeving voor mezelf gemaakt waarin ik alvast ben gaan denken over de tools die ik wil gebruiken voor het webscrapen en gaan nadenken over een stukje code die ik ga schrijven om automatisch de beginnende stappen te doen voor toekomstige websites. Dus ik wil een stuk code schrijven voor dus de basis nu (zoetermeerwijzer) waarin dan word gekeken naar de “Robots.txt” en dan wil ik dat het een stuk terugstuurt over de voorwaardes van de website. In de robots.txt staat namelijk wat de website wel en niet toestaat aan webscraping, waar je wel en niet aan mag komen op de website etc. En dat ik misschien iets kan schrijven dat alle niet allowed url’s of plekken op de website automatisch overslaat of niet meeneemt zodat ik zeker weet dat ik in het proces niet in juridische problemen kan komen door het gebruiken van de informatie waar ik niet bij mag komen. Dit is echter nog een concept dit is niet vandaag af, ik heb ook ondervonden dat de maandag lastig is om te concentreren (en wakker te blijven haha) omdat ik zaterdag en zondag hard werk in de plus en daarna op maandag gelijk weer hier aan de slag moet. Oke dat is genoeg gekletst volgende dag

**Dinsdag 10 September:**

Wederom OV staking, erg vervelend dus ik ben er weer pas om 12:00. Ik heb vandaag een beginnend webscraping script geschreven voor zoetermeerwijzer.nl. Juridisch mag ik dus webscrapen maar ik moet wel letten op het een en anderen. Zo heb je bijvoorbeeld een crawl-delay van 30 seconden dit houd in dat je een 30 seconden delay moet hebben tussen de informatie die je ophaalt (dus de informatie die ik scrape) om ervoor te zorgen de hun website niet overbelast raakt. En ik moet rekening houden met een aantal url’s waar ik niet bij mag komen zoals: "/cms/", "/mijn-profiel/", "/productvergelijking”. Hierbij heb ik dus besloten om alvast te beginnen met een start maken aan scrapen. Dit zorgt ervoor dat ik sowieso wat meer de bal ga laten rollen waardoor ik wat meer dingen kan gaan doen op een dag. En zorgt ervoor dat ik alvast kan gaan experimenteren met scrapen. We hebben het wel eens gedaan op school maar niet bijzonder vaak dus dit word ook weer een beetje schakelen en leren voor mij. Tip voor mezelf, stuur zoetermeerwijzer nog een mail over het plan die we hebben en wat we willen doen met de informatie die we scrapen, hoewel in het robot.txt staat dat we mogen scrapen is het altijd beter om gewoon in direct contact te komen met het bedrijf zodat je weet hoe ze zich voelen bij het idee en mogelijk nog afspraken kan maken over het beheren van hun informatie.

(Ik heb het in de takenlijst gezet)

**Woensdag 11 September:**

Een dagje niets gedaan aan het project, ik moest naar dirksland met een afspraakje en medicijnen ophalen in de middag

**Donderdag 12 September:**

We hebben een soort “Breakthrough” In de opdracht. We hebben echt wel een mooi beginnetje gemaakt aan het webscraping. Ik heb besloten de tool te gebruiken “Scrapy” sinds dit online werd aangeraden om te gebruiken bij grote scraping taken, dit heeft een paar handige tools om bijvoorbeeld automatisch ethisch rekening te houden met de website sinds je een setting heb die er voor zorgt dat alles binnen de robots.txt word uitgesloten bij je gebruik. Ik was vandaag begonnen met het scrapen van <https://zoetermeerwijzer.nl> maar deze agenda van activiteiten blijkt niet binnen hun scraping te liggen, als we verder willen moeten we het bedrijf gaan mailen voor toestemming. Hierdoor ben ik begonnen met de website <https://www.inzet-indewijk.nl/>. Hierbij is het gelukt om ervoor te zorgen dat elke keer als het script runt “soort van realtime” alle evenementen hun basis informatie te scrapen, het is me alleen nog niet gelukt om de activiteiten pagina OOK te scrapen waarin een kleine beschrijving staat van de activiteit. Maar we weten hier nu wel al 2 dingen. We hebben een optie om redelijk simpel de evenementen geupdate te scrapen, en we kunnen ervoor zorgen dat als ik deze website volledig onder de knie heb de code een soort van te recycelen voor websites to come en ik heb een extra excel pagina aangemaakt om bij te houden welke websites ik al heb geprobeerd en bij welke websites ik wel of geen toegang heb om te scrapen, nu moeten we dus ook nog een mailtje gaan maken die we naar de bedrijven kunnen sturen die geen toestemming geven op scraping. Ik heb veel geleerd van scrapy doormiddel van deze bron dus gebruik deze morgen ook maar. PS je bent bij PART 5

<https://www.youtube.com/watch?v=mBoX_JCKZTE>