



Webscraper

Reflectie

**Bachelor in de Toegepaste informatica
keuzerichting Application development**

Wouter Vandueren

Academiejaar 2022-2023

Campus Geel, Kleinhoefstraat 4, BE-2440 Geel

Table of Contents

1.1	Inleiding	2
1.2	Reflectie stageproject	2
1.2.1	Realisatie	2
1.2.2	Voortgang en gebruik project	2
1.2.3	Adviezen opdrachtgever	3
1.3	Persoonlijke reflectie	3
1.3.1	Persoonlijke betekenis stage.....	3
1.3.2	Persoonlijke groei doorheen stage	4
1.3.3	Problemen doorheen stage	4

1.1 Inleiding

In dit reflectiedocument wil ik graag mijn ervaringen en leerpunten delen tijdens mijn stage bij Wisemen. Tijdens deze stage heb ik gewerkt aan het ontwikkelen van een webscraper voor vakantieparken. Dit reflectiedocument biedt mij de mogelijkheid om terug te kijken op mijn stageperiode en te evalueren wat ik heb geleerd en bereikt.

Ik begin dit document met een reflectie over het stageproject dat ik gekregen heb. Hier ga ik eerst over het realisatie proces van het project en wat dit oplevert voor mijn stagebedrijf en de klant. Gevolgd door de verdere voortgang van het project en hoe dit nu al gebruikt kan worden en ik eindig met wat adviezen voor de desbetreffende opdrachtgever.

Vervolgens ga ik over naar mijn persoonlijke reflectie waar ik mijn persoonlijke betekenis van mijn stage beschrijf. Ik ga ook nog beschrijven in dit document op welke punten ik gegroeid ben door mijn stage en welke problemen ik ondervonden heb. Voor deze problemen leg ik dan ook uit hoe ik hiermee ben omgegaan.

1.2 Reflectie stageproject

1.2.1 Realisatie

Tijdens mijn stageperiode heb ik een webscraper gemaakt geschreven in python met behulp van de tool selenium voor de websites van Center Parcs en Belvilla, deze webscraper kan bij Center Parcs bungalows scrapen op basis van vakanties/aanbiedingen, een specifiek land of vakantiepark. Van deze bungalows worden naast de prijs ook andere specificaties gescrept zoals de oppervlakte, het aantal personen, vertrek en aankomstdatum, vakantiepark, land, nieuw prijs, oude prijs, korting in euro's en procent, titel, aantal slaapkamers en de exacte tijd waarop deze bungalow gescrept is. Deze data wordt weggeschreven naar een csv bestand met een specifieke naam en naar een PostgreSQL database in de tabel afhankelijk van wat de user wilt scrapen (cottages, offers of vacations).

Voor deze webscraper is er ook met behulp van Flask een webapplicatie geschreven zodat het voor de user gemakkelijker is om de specificaties waarop de webscraper bungalows gaat moeten scrapen in te geven. De user heeft dezelfde keuzes om bungalows te scrapen als in het gewone python programma maar nu wordt het csv bestand na afloop gedownload op de computer van de user zelf.

1.2.2 Voortgang en gebruik project

De basisvereisten van mijn stageproject zijn afgerond maar dit betekent niet dat het hele project afgerond is, het was mijn opdracht om een voorbeeld te maken van hoe de prijzen van verschillende bungalows van concurrenten van Molenheide gescrept kunnen worden. Uiteindelijk heb ik een goedwerkend voorbeeld gemaakt maar deze is nog niet geïmplementeerd op het portaal van

de eindgebruiker, deze staat nu nog gewoon op mijn github. Normaal gezien was het de bedoeling dat de persoon die backend doet op mijn stage de code ging nakijken die ik geschreven had maar dit is uiteindelijk niet gelukt door tijdsnood langs hun kant. Ik heb wel kunnen tonen dat de webscraper doet wat de klant vraag en deze csv's met de desbetreffende data zijn ook doorgestuurd naar de klant en hier waren ze al heel gelukkig mee. Nu zijn mensen die ook met dezelfde klant bezig zijn op mijn stage met andere dingen bezig zoals de frontend en backend van de website maar uiteindelijk zullen ze mijn voorbeeld webscraper gaan gebruiken of verder afmaken.

1.2.3 Adviezen opdrachtgever

Normaal gezien was het de bedoeling dat er een andere stagiaire samen met mijn ging werken aan deze stageopdracht maar ze hebben uiteindelijk niemand geschikt gevonden. Uiteindelijk was dit geen probleem want ik heb de opdracht toch goed afgerond maar anders was de opdracht misschien nog uitgebreider afgerond dan dat hij nu is. Het was ook leerrijk geweest als iemand van de backend op mijn stageplek mijn code had kunnen nakijken zodat eventuele fouten eruit konden gehaald worden maar dit is uiteindelijk niet gelukt door onvoorzien tijdsnood.

1.3 Persoonlijke reflectie

1.3.1 Persoonlijke betekenis stage

Op mijn stage heb ik veel bijgeleerd over hoe een bedrijf in functioneert binnen de IT

1. Praktische ervaring: Tijdens mijn stage heb ik de mogelijkheid gekregen om praktische ervaring op te doen in het werken binnen een IT-omgeving. Ik heb geleerd hoe bepaalde technologieën en systemen in de praktijk worden toegepast en hoe IT-processen worden uitgevoerd.
2. Technische vaardigheden: Tijdens mijn IT-stage heb ik mijn technische vaardigheden kunnen ontwikkelen en verfijnen. Dit omvat het programmeren in python, werken met databases, flask, selenium. Ik heb ook kunnen meekijken met relevante tools, software en hardware die in de IT-sector worden gebruikt.
3. Probleemoplossend vermogen: Tijdens mijn stage bij Wisemen ben ik ook in staat gesteld om mijn probleemoplossend vermogen te ontwikkelen. Ik werd geconfronteerd met diverse technische uitdagingen en ik had de kans om deze op te lossen met behulp van mijn analytische en kritische denkvaardigheden. Dit bedraagt het oplossen van fouten in mijn code, het uitzoeken hoe ik het best mijn webscraper kan ontwikkelen rekening houdende met de vereisten van de klant.

4. Samenwerking en communicatie: Binnen een IT-bedrijf werk je vaak in teamverband. Een stage geeft je de mogelijkheid om te leren samenwerken met andere IT-professionals, projectmanagers en klanten. Je leert effectief te communiceren, zowel mondeling als schriftelijk, en leert hoe je technische concepten begrijpelijk kunt uitleggen aan verschillende belanghebbenden. Tijdens mijn stage heb ik niet veel hulp gehad van mede developers maar heb ik wel veel meeting gehad met de projectmanager en klant.

1.3.2 Persoonlijke groei doorheen stage

Gedurende mijn stageperiode heb ik actief gewerkt aan het behalen van mijn leerdoelen. Ik heb een functionele webscraper ontwikkeld voor Center parcs en Molenheide, die in staat is om informatie van verschillende pagina's te extraheren en op te slaan in een gestructureerd formaat. Door dit project heb ik mijn programmeervaardigheden aanzienlijk verbeterd, vooral op het gebied van web scraping en het werken met externe bibliotheken.

Tijdens het bouwen van de webscraper heb ik veel geleerd over het proces van data scraping, data cleaning en dataopslag. Ik heb verschillende technieken toegepast om relevante informatie uit de HTML-structuur van de websites te halen en deze gegevens op een consistente manier opgeslagen. Ik heb ook geleerd om te gaan met uitdagingen zoals anti-scrapingmaatregelen en het omgaan met ongestructureerde data.

Het ontwikkelen van efficiënte en onderhoudsvriendelijke code was een belangrijk aspect van mijn stageproject. Ik heb geleerd om code te schrijven die modulair is, goed gedocumenteerd en gemakkelijk te begrijpen voor andere ontwikkelaars. Door het werken met version control, zoals Git, heb ik ervaring opgedaan met het werken in een team en het beheren van wijzigingen in de codebase.

Daarnaast heeft mijn stageperiode mijn kennis van webtechnologieën vergroot. Ik heb een dieper inzicht gekregen in HTML-structuur, CSS-selectoren en DOM-manipulatie.

1.3.3 Problemen doorheen stage

- Doordat het het bedrijf waar ik stage doe enorm gegroeid is op vlak van personeel is het internet niet meer sterk genoeg als er veel mensen op kantoor zijn. Het probleem hiermee was dat mijn webscraper stopte met werken als hij geen nieuwe bungalows zag op de pagina, mits het internet zo slecht was konden deze niet op tijd ingeladen worden. Dit probleem kon ik verhelpen door de time.sleeps iets langer in te stellen maar dit zorgde er wel voor dat het programma veel langer duurde als normaal.
- Doordat ik in een redelijk drukke periode mijn stage begon kon ik niet echt op veel hulp rekenen van de mensen die in mijn tribe werkten. Er was maar 1 senior backender in de tribe en die had het te druk met zijn werk en het opleiden van junior backenders. Dit vormde gelukkig neit een al te groot probleem want ik heb het project toch afgekregen. Maar als iemand mijn code had kunnen nakijken dan was deze nog beter geweest.